In compliance with the
Canadian Privacy Legislation
some supporting forms
may have been removed from
this dissertation.

While these forms may be included
in the document page count,
their removal does not represent
any loss of content from the dissertation.

University of Alberta

Characterization of the UNC-119 Family in Mammals

by

Chantal Denholm

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Molecular Biology and Genetics

Department of Biological Sciences

Edmonton, Alberta

Fall, 2003

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisisitons et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-87748-5
Our file Notre référence
ISBN: 0-612-87748-5

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

Canadä

University of Alberta

Library Release Form

Name of Author: Chantal Denholm

Title of Thesis: Characterization of the UNC-119 Family in Mammals

Degree: Master of Science

Year this Degree Granted: 2003

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

August 8, 2003

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Characterization of the UNC-119 Family in Mammals" submitted by Chantal Denholm in partial fulfillment of the requirements for the degree of Master of Science in Molecular Biology and Genetics.

Aug 8/03

Abstract

The study of highly conserved proteins and their cellular functions is critical for the characterization of common processes underlying development. This approach has been adopted to determine the biological function and to delineate the evolution of the highly conserved UNC-119 protein family. Previous studies point to extensive functional conservation of UNC-119 across metazoans despite vast differences in protein distribution. Invertebrate UNC-119 is expressed pan-neurally throughout life, while human UNC119 has been localized to the adult retina. Characterization of the mammalian UNC-119 branch reveals that UNC119 is widely expressed during development and is found non-retinally in adults. A second highly related homologue, UNC119B, has been identified that is also widely temporally and spatially expressed in retinal and non-retinal tissues. Preliminary analysis of their structural and biochemical properties suggests that UNC-119 proteins are universal receptor-associated adapter molecules that play a critical role in signal transduction via activation of Src tyrosine kinase cascades.

Dedication

I would like to dedicate this manuscript to my mother, Madeleine Denholm. She has been my primary source of inspiration and support throughout my life and in particular during my completion of this project.

Without her continued encouragement, support, help, patience, and interest, the following manuscript would be but a fraction of what is presented to you here. Merci beaucoup Maman!

Acknowledgments

I would like to thank the National Sciences and Engineering Council of Canada, the Alberta Heritage Foundation for Medical Research, the NeuroScience Canada Foundation, and the University of Alberta for their generous funding and learning opportunities.

I would also like to extend my most sincere gratitude towards my committee members, Drs. HE McDermid and CT Putman for their tremendous support. Both have been instrumental to the completion of this project through many levels. They have generously devoted time, protocols, reagents, lab space, equipment, and technical expertise that were critical to the completion of my experiments.

I would also like to thank R Bhatnagar, R Mandryk, and J Scott who have kindly assisted with documentation and photography of the results presented here. They have also provided me with resources and histological training over the course of the project as well as continued friendliness and emotional support.

I thank Drs. K Bagnall and S Zalik for sharing with me their histological expertise and assistance in interpretation of experimental results.

Finally, I would also like to thank my supervisor, Dr D Pilgrim for accepting me into his laboratory and for his continued supervision and guidance throughout this project!

Table of Contents

Introduction 1

1.1 Why is the study of development important? 1

1.2 *Caenorhabditis elegans unc-119* 3

1.3 UNC-119 is highly conserved across metazoans 5

1.4 UNC-119 expression patterns are similar across invertebrates 6

1.5 Differential expression and UNC-119 branching in fish 7

1.6 Mammalian UNC-119: Restricted to the adult retina? 8

1.7 Exploring a dichotomy: Reason for further UNC-119 characterization in mammals 10

1.8 Project objectives and approaches taken 15

Materials and Methods 17

2.1 Northern analysis 17

2.2 Mouse embryo collection and storage 17

2.3 Whole mount *in situ* hybridization 18

2.4 *In situ* hybridization of thin tissue sections 19

2.5 Antibody production 20

2.6 Antiserum purification 21

2.7 Western analysis 21

2.8 Mammalian tissue collection for protein extraction 21

2.9 Non-denaturing protein extraction 22

2.10 Immuno-chemistry 22

2.11 Cell culture 23

2.12 Tagged-protein pull-down assay **23**

2.13 Antibody immuno-precipitation assay **24**

Results **25**

3.1 UNC119B: A second UNC119 homologue in mammals **25**

3.2 UNC119 and UNC119B are widely expressed during development **31**

3.3 Whole mount *in situ* hybridization shows ubiquitous expression **32**

3.4 *In situ* hybridization of mouse embryo sagittal sections **33**

3.5 Production of α-UNC119 and α-UNC119B antibodies **36**

A) Expression of GST-fusion proteins in a bacterial system **36**

B) Purification of full-length GST-tagged UNC119 and UNC119B proteins **37**

C) Antibody specificity against recombinant protein **37**

D) Antibody specificity against retinal protein extracts **37**

E) Antibody examination against brain extracts **39**

F) Antibody examination against mouse extracts **39**

3.6 UNC119 and UNC119B protein localization in mouse embryos **42**

3.7 PC-12 cell culture preparation **43**

3.8 Identification of UNC119 interacting proteins **44**

Discussion **46**

4.1 UNC119B: A second mammalian UNC-119 homologue **46**

4.2 Vertebrate UNC-119 expression patterns **52**

4.3 Distribution of mammalian UNC-119 proteins **55**

4.4 UNC119 interacting proteins **58**

4.5 Vertebrate UNC-119 family: General conclusions and speculations **60**

Concluding Remarks **65**

Tables **66**

Figures **69**

Bibliography **147**

Appendices **155**

List of Tables

1. Potential UNC-119 homologues in Fugu. **66**

2. Densitometry analysis of α-UNC119 signal intensity on Western blots. **67**

3. Summary of the migration rates of the products detected by α-UNC119 and α-UNC119B antisera. **68**

List of Figures

1. *C. elegans unc-119* mutants have a disorganized nervous system. **69**
2. UNC-119 is present in axons and cell bodies of *C. elegans* neurons. **71**
3. A phylogenetic tree of the UNC-119 protein family. **73**
4. A comparison of the amino acid sequences of UNC-119 proteins in different organisms. **75**
5. DmUNC-119 expression during embryogenesis in *Drosophila melanogaster*. **77**
6. *Danio rerio unc119* and *unc119b* expression and protein localization. **79**
7. Mammalian UNC119 and ARL2 co-localize in the adult retina. **81**
8. Amino acids that may mediate the UNC119-ARL2 interaction are conserved across UNC-119 proteins. **83**
9. Human UNC119B: A homologue of UNC119. **85**
10. The structure of human UNC119B. **87**
11. Mammalian UNC119B. **89**
12. The genomic location of mouse UNC119B. **91**
13. The gene structure of mouse UNC119B. **93**
14. UNC119 and UNC119B are widely expressed in the human fetus. **95**
15. UNC119 and UNC119B mRNA expression in mouse E10.5 embryos. **97**
16. UNC119 and UNC119B mRNA expression in mouse E11.5 embryos. **99**
17. UNC119 and UNC119B mRNA expression in mouse E12.5 embryos. **101**
18. UNC119 *in situ* hybridization of mouse embryonic day 8, 10 and 14 sections. **103**
19. UNC119B *in situ* hybridization of mouse embryonic day 8, 10 and 14 sections. **105**

20. Induction of UNC119-GST and UNC119B-GST fusion proteins in bacterial cells. **107**
21. Purification of GST-fusion proteins from the bacterial strain BL21(DE3) following induction. **109**
22. Western blot of α-UNC119 and α-UNC119B antisera against purified recombinant protein. **111**
23. Western blot of α-UNC119 antisera against bovine retinal protein. **113**
24. Western blot of purified α-UNC119B against bovine retinal protein. **115**
25. Characterization of purified α-UNC119B. **117**
26. Western blot of α-UNC119 antisera against adult bovine brain protein. **119**
27. Western blot of α-UNC119 antisera against mouse embryo extracts. **121**
28. Western blot and densitometry analysis of α-UNC119 purified from OF3. **123**
29. Characterization of products recognized by α–UNC119 purified from OF3. **125**
30. Western blot of α-UNC119B against mouse embryo protein. **127**
31. Western blot of α–UNC119B immuno-precipitates. **129**
32. Un-purified α-UNC119 immuno-localization on mouse cryo-sections. **131**
33. Immuno-localization of UNC119 on mouse embryo sections. **133**
34. Immuno-localization of UNC119B on mouse embryo sections. **135**
35. PC-12 cells produce UNC119 and UNC119B. **137**
36. Protein gel of UNC119-GST binding proteins. **139**
37. Protein gel of mouse embryo proteins that immuno-precipitate with α-UNC119. **141**
38. The UNC-119 protein family. **143**
39. A wild speculation of UNC-119 function. **145**

List of Appendices

1. Detailed protocol for whole mount *in situ* hybridization. **155**

2. Detailed protocol for *in situ* hybridization of sections. **158**

3. Nucleotide sequences of *in situ* hybrization probes. **160**

4. Detailed protocol for immuno-histochemical staining. **161**

5. Peptide summary report for a α–UNC119B immuno-precipitate. **162**

6. α-UNC119 co-immuno-precipitate protein summary report: Ig Heavy Chain. **164**

7. α-UNC119 co-immuno-precipitate protein summary report: A-X actin. **165**

8. α-UNC119 co-immuno-precipitate protein summary report: EpoR. **167**

List of Abbreviations

Ab	Antibody
ATCC	American Type Culture Collection
BCIP	5'-Bromo-4-chloro-3-indolyl-phosphate
bp	Base Pair
BSA	Bovine Serum Albumin
DAB	3,3'-Diaminobenzidine
DEPC	Diethyl Pyrocarbonate
DIG	Digoxigenin
EDTA	Ethylenediamine Tetra-Acetic Acid
EGTA	Ethylene Glycol-bis-(b-amino-ethyl ether) N,N,N′,N′-Tetra-Acetic Acid
EST	Expressed Sequence Tag
GABA	Gamma-Aminobutyric Acid
GCL	Ganglion Cell Layer
GFP	Green Fluorescent Protein
GST	Glutathione S-Transferase
hpf	Hours Post Fertilization
IgG	Immunoglobin
INL	Inner Nuclear Layer
IPL	Inner Plexiform Layer
IS	Inner Segment
ISH	*In Situ* Hybridization
Kbp	Kilobase
KDa	Kilodalton
MTN	Multiple Tissue Northern
NBT	Nitroblue Tetrazolium Chloride
o/n	Overnight
ONL	Outer Nuclear Layer
OPL	Outer Plexiform Layer
ORF	Open Reading Frame
OS	Outer Segment
PAGE	Poly Acrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PBT	Phosphate Buffered Saline + 0.1% Tween 20
PEST	Sequences enriched in Proline (P), Glutamic acid (E), Serine (S), and Threonine (T)
PI	Pre-Immune
PMSF	Phenylmethylsulfonyl Fluoride
RPE	Retinal Pigment Epithelium
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SAGE	Serial Analysis of Gene Expression
SDS	Sodium Dodecyl Sulfate
SH	Src Homology Domain
SSC	Saline Sodium Citrate
TBS	Tris Buffered Saline

TBT	Tris Buffered Saline + 0.1% Tween 20
TCA	Trichloroacetic Acid
TE	Tris-EDTA
TLCK	L-1-Chloro-3-(4-tosylamido)-7-amino-2-heptanone hydrochloride N-alpha-Tosyl-L-lysine chloromethyl ketone
URL	Uniform Resource Locator
VNS	Ventral Nervous System
WMISH	Whole Mount *In Situ* Hybridization

Introduction

The purpose of this M.Sc. thesis is to characterize the biological roles and evolution of the highly conserved UNC-119 protein family within a developmental context. In the following pages, the importance of developmental biology and of the study of highly conserved gene families will be discussed, followed by a review of what is known about UNC-119 in several model systems. Finally, remaining questions regarding UNC-119 that this project aims to address and the specific objectives of the study will be outlined.

1.1 Why is the study of development important?

Even though our world is populated by a tremendous variety of life forms, there are many common means by which the development of different organisms is achieved. Development is the essential process of building a living multi-cellular organism from a single cell. This intricate process is incredibly complex, because in addition to the generation of multitudes of cell types and their plastic organization into functioning tissues and organ systems, every animal must function as it builds itself (Gilbert 2000). Furthermore, an understanding of this process is essential to understanding many other areas of biology due to its underlying basis linking anatomy, physiology, molecular biology, oncology, immunology, and ecology with evolution.

In many cases, vast differences in morphology, structure, or behaviour between organisms stem from developmental differences that occur during later embryological stages. Many underlying processes remain remarkably unchanged. Even across entire Kingdoms of species such as the heterotrophic metazoans, several genes with homologous structures and functions have been identified. As expected, these genes tend to have regulatory functions in the fundamental mechanisms controlling cell and tissue interactions during early development (Kaufman 1995). Tightly regulated developmental processes include cell fate specification and differentiation, pattern formation, signal transduction, and programmed cell death (Gilbert 2000).

For many of these genes, their conserved properties across organisms has been examined to better understand their biological functions. Several studies of the most highly conserved genes have even led to the dissection of mechanisms fundamental to basic general development across species. For example, the *Distal-less/DLX* genes found in numerous phyla of protosomes and deuterosomes are involved in appendage formation (Lee and Jacobs 1999). The *Bcl2* genes found in nematodes and vertebrates are involved in programmed cell death (Vaux *et al.* 1992), and the common metazoan *lin-12/Notch* receptor/ligand pairs are critical to proper cell fate specification (Greenwald 1985).

Furthermore, functional conservation across metazoans for several genes has also been clearly demonstrated using cross-species ectopic expression experiments. For example, Halder and colleagues (1995) induced ectopic formation of eyes in fly using the mouse *Pax6* gene. These fully differentiated eyes were morphologically normal and were inducible in several different fly tissues depending on the primordial cells targeted to express the mouse gene. Repetition of Halder's experimental design has since led to the identification of several other conserved genes (*sine oculis, eyes absent, dachshund*) involved in eye formation (Jean *et al.* 1998; Relaix and Buckingham 1999). Such conservation is remarkable given the extensive structural differences between the compound fly eye and single lens mouse eyes.

One of the extraordinarily conserved yet poorly understood fundamental processes of developmental biology is neurogenesis. The nervous system plays a critical role in communication, coordinating responses between cells, tissues, and systems in multi-cellular organisms. Its extensive conservation is a virtue of this role, because in addition to becoming specified to adopt a neuronal fate, neural cells must grow extensively, migrate, and form extremely specific interconnections to generate a functioning system that links the many different tissues of the body to generate a viable organism.

Given the complexity and importance of this process, its dissection at the molecular level is critical to understanding the mechanisms underlying basic development. Studies of the most highly conserved genes involved in this process are underway, and several important protein families have been identified. Fortunately, their remarkable conservation can be exploited for study, because the differing advantages of

several model organisms may be utilized and applied in comparative analyses to determine their biological roles. This is the approach that has been adopted to study the role of the unc-119 family in development of the nervous system.

1.2 *Caenorhabditis elegans unc-119*

Unc-119 is an exceptionally conserved gene family. It was initially identified and cloned in the nematode *Caenorhabditis elegans* because of its requirement for proper development and maintenance of this organism's nervous system. *Unc-119* mutants display a variety of phenotypic traits that are characteristic of neurological defects. Similar to other nematode *unc* gene mutants, *unc-119* animals display a recessive uncoordinated phenotype and reduced mobility. They are also defective in several neuron-mediated processes resulting in sensory and behavioural abnormalities. They exhibit constitutive pharyngeal pumping (a feeding behavior) even in circumstances of food depletion, and they are unable to enter the dauer state, a developmentally arrested morphological state in response to overcrowding. Another neuron mediated behaviour, egg laying, is also defective in *unc-119* mutants, because hermaphrodites retain an elevated number of their embryos in their uteri due to an inability to correctly lay eggs (Maduro and Pilgrim 1995).

Fine structural analysis of *unc-119* mutants demonstrates that their gross phenotype is attributable to a disorganized nervous system. As demonstrated in figure 1, neurons in *unc-119* worms are excessively branched, and their axon bundling, spacing, and positioning are disrupted. Separation (defasciculation) of individual motor neuron and inter neuron axons from the dorsal and ventral nerve bundles is prominent, and spacing of the commissural neurons that are normally extended along the circumference of the worm at regular intervals is also severely disrupted. Abnormal and excessive branches of *unc-119* commissures are frequent, and they tend to initiate from regions where physical boundaries are encountered during axon outgrowth (Materi, personal communication). However, even though every *unc-119* mutant exhibits all of these structural defects, specific neurons are affected in variable patterns in individual animals (Materi, personal communication).

These structural defects are consistent with the observed *unc-119* phenotype. Proper motor neuron positioning and formation of neuromuscular junctions is critical to mobility and coordination, and although an uncoordinated phenotype could also result from muscular defects, muscle ultrastructure, sarcomere placement, and acetylcholine receptors are normal in *unc-119* worms (Maduro and Pilgrim 1995). Furthermore, many *unc-119*-specific behaviours may also be accounted for by the structural defects observed in mutant worms. For example, the ASI amphid neurons, essential for dauer formation, are excessively branched and improperly connected to the nerve ring in *unc-119* worms (Materi, personal communication). These worms are dauer defective.

Expression and protein localization studies demonstrate that *unc-119* appears to be expressed pan-neurally throughout life (Maduro and Pilgrim 1995). Expression has even been observed in embryonic stages as early as the 60 cell stage (Materi, personal communication), before neural differentiation has taken place (Maduro *et al.* 2000). As differentiation occurs, *unc-119* is expressed in a cell autonomous manner throughout neuronal cells. As shown in figure 2, the UNC-119 protein is also present in the cells and is concentrated predominantly along their projecting axons and to a lesser extent in their cell bodies (Knobel *et al.* 2001).

Temporal analysis of mutants reveals that development of the nervous system may involve several independent processes. Time-lapse video imaging of commissure formation reveals normal branching during axonal extension in *unc-119* worms albeit at a reduced rate of migration. It is only subsequent to initial outgrowth that projected axons undergo retraction accompanied by sprouting of secondary growth cones and branches (Knobel *et al.* 2001). Rescue experiments demonstrate that the *unc-119* branching phenotype is completely suppressible by post developmental protein expression. Furthermore, reversal of pre-existing branching defects is even observed in developing mutants following rescue, indicating that UNC-119 is dispensable for initial establishment of neuronal branching. The protein appears to function in maintenance of cell polarity and stability of the nervous system.

However, even though UNC-119 is not required for initial outgrowth of un-branched axons, proper positioning of these axons is dependent on UNC-119 throughout their development. During the early migration phase of neuronal development, axon

bundling and cell body placement defects are observed in *unc-119* worms even though branching appears normal (Materi, personal communication).

This suggests that cell placement, fasciculation and growth cone targeting may be independent processes, and that UNC-119 may have multiple roles in development of the nervous system. These could include a stabilizing function in branching and synapse distribution via cytoskeletal (microtubule) maintenance, or via membrane trafficking and vesicle targeting regulation (Knobel *et al.* 2001). A second independent role for the protein could be activation of proteins involved in fascicle formation and stability (Materi, personal communication).

1.3 *UNC-119* is highly conserved across metazoans

Although the *unc-119* mutant phenotype is well characterized in nematodes, its biochemical role within the cell remains to be determined. The *unc-119* DNA and protein sequences bear little similarity to other genes, motifs, or domains of known function and therefore do not reveal clues about its biochemical properties. However, even though *unc-119* is not homologous to other genes of known function, it is very highly conserved across metazoans.

Functional *unc-119* homologues have been identified in other invertebrates (*Drosophila melanogaster*), and also in many vertebrates such as fish (*Danio rerio*) and mammals, as demonstrated in figure 3. There is extensive unc-119 conservation at both the genomic and protein levels, suggesting that they likely evolved from a common ancestral form. All share a common intron in the most highly conserved region of the gene (Maduro *et al.* 2000), and the UNC-119 proteins are also very similar, sharing greater than 50% identity across phyla. The human and nematode proteins, for example, are 57% identical (Higashide *et al.* 1998). Between more closely related species, such as across mammals or between nematodes, UNC-119 identity exceeds 90% (Higashide *et al.* 1998; Swanson *et al.* 1998; Maduro and Pilgrim 1996).

As demonstrated in figure 4, the regions of identity between the proteins are not evenly distributed, suggesting a bipartite structure with 2 conserved domains. The amino terminal of the proteins is very poorly conserved even though the carboxyl half is very

highly conserved, sharing greater than 70% identity across species. This suggests that it likely encodes the active or catalytic domain of UNC-119 (Maduro *et al.* 2000). Another moderately conserved domain is also present in the latter portion of the amino terminal, therefore it may have functional properties (Maduro *et al.* 2000). Supporting the extensive structural conservation observed between the UNC-119 proteins, transgenic experiments in mutant worms using both invertebrate and vertebrate UNC-119 counterparts demonstrate that conservation of protein function is maintained across metazoans and requires only the carboxyl domain, confirming its biochemical importance (Maduro *et al.* 2000). UNC-119 proteins are also distantly related to the PDEδ/PDL-1 family, and, as described in detail later, the known crystal structure of PDEδ could potentially be informative in analysis of the UNC-119 functional domain.

1.4 UNC-119 expression patterns are similar across invertebrates

Given this high level of conservation at the structural and functional level, further insight into the biological roles of UNC-119 can be derived by comparing its expression pattern and protein distribution across phyla. In all invertebrates examined to date, unc-119 transcripts are of relatively low abundance in expressing tissues (Maduro *et al.* 2000).

In the fruit fly *Drosophila melanogaster*, DmUNC-119 is expressed in a pattern similar to that observed in nematodes. As demonstrated in figure 5, it is first detected embryonically after the first neurons begin to differentiate from neuroblast precursor cells. Consistent with observations in different *Caenorhabditis* species (Maduro and Pilgrim 1996), it continues to be expressed in all neurons throughout life in the nervous system, and protein translation is detectable in embryonic, larval, and adult stages (Hawley 2002).

A putative null DmUNC-119 mutant (*attila*) has been identified. *Attila* embryos have a disorganized nervous system that has defects consistent with those observed in nematodes. Unlike in worms, however, *attila* displays a late recessive embryonic lethal phenotype (Hawley 2002). Complementary genetic analysis reveals a synergistic phenotype when a DmUNC-119 deficiency is crossed to a *chickadee* mutant background,

suggesting that the proteins may function in the same pathway (Hawley 2002). *Chickadee* is a conserved homologue of Profilin, a key regulator of actin filament assembly involved in the actin polymerization required for specific aspects of growth cone behaviour (Wills *et al.* 1999).

1.5 Differential expression and UNC-119 branching in fish

Although structural and functional conservation of UNC-119 is maintained across metazoans, the vertebrate temporal and spatial patterns of unc-119 expression are remarkably different from those observed in invertebrates. In the zebrafish (*Danio rerio*), a simple and easily manipulated system, multiple unc-119 homologues have been identified and encode at least 2 non-redundant functional proteins: Unc119 and Unc119b. However, their expression profiles, while similar temporally, transcend the nervous system and result in non-overlapping sites of protein localization (Manning 1999).

Unc119 and *unc119b* are both expressed in the early embryo (when zygotic transcription begins) and continue through to adulthood. Similar to invertebrates, expression is first detected before overt patterning of the nervous system or axon outgrowth has taken place. However, even though expression of both homologues is present in adults (as found in invertebrates), adult expression appears to show sexual differences and circadian fluctuations (Manning 1999).

The spatial pattern of *unc119* expression (shown in figure 6A) appears to be ubiquitous throughout embryonic and larval development. Western analysis indicates that its product is first detected at 4 hours post fertilization (hpf) and continues to be present well into the swimming larval stages, the oldest age examined. As demonstrated in figures 6B and 6C, Unc119 protein localizes to many neural structures, including a specific photoreceptor cell layer of the differentiating retina, the post-optic commissure, the nasal pits, the neuromast cells in the trunk, and certain neuronal cells in the fins (Manning, personal communication). Knockdown of Unc119 function using an antisense morpholino approach (Nasevicius and Ekker 2000) results in a specific curly tail down phenotype, characterized by defects including a shortened body axis, an altered hindbrain, reduced circulation and response to touch, and increased numbers and

branching of neurons (Manning, personal communication). Such features are consistent with the unc-119 phenotype in invertebrates. In fact, when separated from their yolks, these curved knockdown embryos look remarkably like *unc-119* worms.

Furthermore, genomic analysis demonstrates that *unc119* and the very closely linked *whnb* (Shorpp *et al.* 2002) map to the zebrafish linkage group that shows conserved synteny with human chromosome 17q11-12, the location of UNC119 and hFOXN1 (the human *whn* orthologue) (Swanson *et al.* 1998; Higashide and Inana 1999; Shorpp *et al.* 2002). UNC119 and Unc119 are therefore likely true orthologues (Manning, personal communication).

Similar to *unc119*, *unc119b* is expressed ubiquitously in embryos (figure 6D) and its encoded protein is detectable throughout embryonic and larval stages. In addition, the Unc119b knockdown phenotype displays, at least superficially, only slight differences from Unc119. Despite such similarities, however, the spatial patterns of Unc119b protein distribution differ sharply from Unc119, and they contrast with those observed in invertebrates. As opposed to a neuronal localization, Unc119b is largely found in several slow twitch muscles throughout the body (figure 6F), including many found in the facial and jaw regions, the trunk, and the ocular muscles. Like Unc119, Unc119b also localizes to photoreceptor cells of the retina, though Unc119 and Unc119b are exclusively restricted to different cell layers of the retina, as depicted in figures 6B and 6E (Manning, personal communication).

1.6 Mammalian UNC-119: Restricted to the adult retina?

Mammalian UNC119 was initially identified as HRG4 (Human Retinal Gene 4), because it was isolated from a screen enriching for retinal specific transcripts (Higashide *et al.* 1996). It has since been renamed UNC119 (Higashide and Inana 1999). As opposed to widespread expression of other unc-119 family members, human UNC119 expression is reported to be limited to adult stages and restricted to the retina, specifically located in the ribbon synapse nerve junction as shown in figure 7A (Higashide *et al.* 1998). Its encoded protein localizes to the pre-synaptic zone of the junction and appears to associate with synaptic vesicles (Higashide *et al.* 1998). A truncation mutation in UNC119 has

been observed in one patient with a late-onset retinal degeneration disease (cone-rod dystrophy) (Kobayashi *et al.* 2000), and data obtained from a corresponding transgenic model suggests that the degenerative phenotype results from a decrease in peripheral membrane proteins of the synaptic vesicle (Kubota *et al.* 2002).

It has recently been shown that UNC119 interacts with the GTP binding factors ADP-ribosylation factor like proteins 1, 2, and 3 (ARL1, ARL2, ARL3) in an activation state independent manner (Van Valkenburgh *et al.* 2001). The UNC119-ARL2 interaction has been confirmed by an independent group, and co-localization of both proteins has been observed in the outer plexiform layer of the retina and in the inner segments of the photoreceptors as shown in figure 7B (Kobayashi *et al.* 2003).

ARL proteins are related to the ADP ribosylation factor (ARF) family of highly conserved, ubiquitous GTP binding proteins that regulate membrane trafficking, organelle structure, and regulation of vesicular transport (Donaldson and Jackson 2000; Hanzal-Bayer *et al.* 2001). Their specific biochemical functions, however, have not been delineated (Kobayashi *et al.* 2003).

Other studies indicate that ARL1, ARL2 and ARL3 also interact with several other proteins, including PDEδ, a protein that is 30% identical to UNC119 in the functional carboxyl domain (Li *et al.* 1998, Van Valkenburgh *et al.* 2001). The PDEδ amino acid residues involved in the ARL2 interaction have been delineated by X-ray crystallography (Hanzal-Bayer *et al.* 2001), and these are highly conserved (77% identical) with UNC119, suggesting that UNC119 may bind ARL2 through the same domain (Kobayashi *et al.* 2003). Furthermore, as would be expected for functionally important residues, these amino acids are also very highly conserved across the entire UNC-119 protein family (69% identical), and they are perfectly conserved in vertebrates, as demonstrated in figure 8.

Like human UNC119, PDEδ was originally identified in retina. It was isolated from a rod-specific phosphodiesterase complex (Gillespie *et al.* 1989) and appears to be a key component of visual transduction cascade in vertebrate photoreceptor cells (Li *et al.* 1998). It solubilizes certain subunits of the phosphodiesterase complex from the photoreceptor membrane by binding their prenylated ends (Florio *et al.* 1996), and it interacts with several proteins (in addition to ARL2) including some derived from known

retinal disease causing genes, such as RPGR (retinitis pigmentosa GTPase regulator) (Linari *et al.* 1999).

However, even though UNC119 is present in the same cell types and interacts with at least one class of PDEδ's binding partners, UNC119 and PDEδ are not redundant, because UNC119 is unable to solubilize proteins from the photoreceptor membrane (Li *et al.* 1998). Their binding dynamics are also distinct, because UNC119 binds ARL proteins independently of their GTP activation state while PDEδ binds them only in the activated, GTP-bound state (Van Valkenburgh *et al.* 2001).

Based on these studies, UNC119 has been postulated to regulate membrane protein trafficking in synaptic vesicles and microtubule formation in photoreceptor axons via ARL proteins (Kobayashi *et al.* 2003). Consistent with this hypothesis, there is precedence for microtubule regulation in the retina by other ARL proteins such as ARL3 (Kobayashi *et al.* 2003). However, when considering the pattern of UNC-119 expression and localization in other species, a more widespread role for the protein is suspected.

1.7 Exploring a dichotomy: Reason for further UNC-119 characterization in mammals

Given the extensive structural and functional conservation of UNC-119 in metazoans, the drastically contrasting mammalian expression pattern compared to other organisms makes it difficult to postulate common biochemical function(s) for this highly conserved protein family. Expression of mammalian UNC119 in the synaptic termini of the rod and cone photoreceptors of the eye is consistent with a neurological role of the protein, however this limited localization would suggest that its function is restricted to visual signal neural transmission (Higashide *et al.* 1998; Swanson *et al.* 1998). This function is unlikely in other organisms such as nematodes that do not have eyes or other photo-transducing appendages!

This discrepancy in localization of mammalian UNC119 could be resolved by the non-mutually exclusive possibilities that either its expression is not restricted to the retina, or that, as found in zebrafish, multiple mammalian UNC-119 homologues exist. The combined expression patterns of several related proteins could be consistent with

invertebrate observations for UNC-119 with the addition of novel, mammalian-specific, specialized sites such as the retina.

There is reason to explore a non-retinal cellular role for UNC-119 in mammals. Although it has independently been reported to be enriched in the retina of other mammalian species such as mouse, rat, and dog (Lin and Sargan 2001), its expression has not been closely examined in other tissues or during developmental stages. There is also indirect genomic database evidence for mammalian UNC119 expression at lower levels in multiple tissues. UNC119 transcripts have been sequenced from tissues including human fetal brain and mouse embryo (Swanson *et al.* 1998), and microarray analysis demonstrates that it is up-regulated in classic medullobalstomas, the most common embryonic malignant brain tumor (Pomeroy *et al.* 2002). UNC119 is therefore likely expressed in the normal developing brain.

Another study developing a full catalog of rod photoreceptor-enriched transcripts using serial analysis of gene expression (SAGE) has confirmed strong retinal enrichment of UNC119 in the differentiated retina. However, the UNC119 expression pattern does not match that of other known rod-specific transcripts. Unlike other photoreceptor-specific genes, UNC119 transcripts are detectable in mouse fibroblasts, and no enrichment of UNC119 in the outer nuclear retinal layer (the location of rod cells) compared to whole retina extracts is observed (Blackshaw *et al.* 2001). UNC119 is also slightly up-regulated in Crx deficent retinas even though rod-enriched gene expression is typically dependent on functional Crx protein (Blackshaw *et al.* 2001). Crx is a transcription factor that is expressed in developing and adult photoreceptors and regulates photoreceptor differentiation (Chen *et al.* 1997 and Furukawa *et al.* 1997). Abundant retinal transcripts whose expression are independent of Crx, or are down-regulated by the protein tend to have a ubiquitous distribution (Blackshaw *et al.* 2001). Finally, Swanson *et al.* (1998) have directly shown faint adult UNC119 expression in adrenal, cerebellum, and kidney tissues using RT-PCR analysis. Non-retinal UNC119 expression is therefore present and remains to be characterized.

Other reasons for exploring a non-retinal role for UNC119 stem from the mammalian retinal studies themselves. The significance of the patient with the UNC119 mutation, and the phenotype of the transgenic mouse carrying the corresponding mutation

are questionable. Only a single heterozygous patient has been identified with the UNC119 mutation, therefore the possibility that this individual's phenotype is the result of an unrelated mutation cannot be excluded. Furthermore, even though the transgenic mouse carrying the mutation displays a retinal phenotype consistent with the one observed in the patient, the transgene used to create the mouse was driven by the rhodopsin promoter as opposed to its native one (Kobayashi *et al.* 2000). Because rhodopsin is a very highly expressed and tightly regulated retinal rod photoreceptor specific gene (Zack *et al.* 1991; Nie *et al.* 1996), using it to drive UNC119 expression creates an artificial situation that may not reflect the effects of an endogenous mutation. Transgenic UNC119 has not been examined nor expressed in non-retinal tissues.

Also, even though ARL2 was isolated as an UNC119 interacting protein using a human retinal cDNA library by one group (Kobayashi *et al.* 2003), ARL2 has also co-precipitated UNC119 from a human B cell library, confirming translation of both proteins in non-retinal cells (Van Valkenburgh *et al.* 2001). Furthermore, expression analysis reveals that ARL2 is ubiquitously distributed throughout embryogenesis and adulthood (Clark *et al.* 1993), attaining its highest levels in neural tissue (Sharer *et al.* 2002). This is a distribution that might be expected for UNC-119 in mammals.

Orthologues of ARL2 in other species have been also been identified, and these GTPases appear to play universal roles in microtubule dynamics as opposed to specific functions in phototransducing retinal cells. The yeast ARL2 ortholog Alp41 is an essential gene required for proper folding of tubulins and their incorporation into microtubules (Radcliffe *et al.* 2000), and a plant ortholog (TTN5) is present throughout the plant lifecycle. It appears to function in cytoskeletal organization and microtubule regulation of seed development (Tzafrir *et al.* 2002). Contrary to the idea that a protein-protein interaction between UNC119 and ARL2 confirms a retinal-specific role for UNC119 (Kobayashi *et al.* 2003), this suggests that both proteins may have a more general function in cytoskeletal regulation. However, because ARL proteins are found outside metazoans unlike (UNC-119 or PDEδ), a more widespread ARL distribution pattern could be required for specific ARL functions that are independent of UNC-119 or PDEδ.

ARL2 interacts with the distant relative of UNC119 or PDEδ (Florio *et al.* 1996). Like UNC119, PDEδ is predominantly transcribed in photoreceptors and has been associated with retinal disease (Li *et al.* 1998). However, further analysis demonstrates that it is also expressed at lower levels in multiple tissues (Marzesco *et al.* 1998), and there is even direct evidence of PDEδ function in other cell types. It solubilizes several non-retinal GTP-associated proteins, such as the GTPase Rab13 involved in membrane trafficking at tight junctions between polarized epithelial cells (Marzesco *et al.* 1998). This distant UNC119 relative is therefore not specific to the rod phosphodiesterase complex, playing a more universal role in signaling cascades (Li *et al.* 1998). This suggests that many retinal enriched proteins actually have widespread functions.

Finally, unc-119 transcripts are of relatively low abundance in other organisms examined to date (Maduro *et al.* 2000), contrasting with the high levels reported in the retina (Higashide *et al.* 1998). The possibility of lower levels of mammalian expression not visible at the sensitivity used during initial screening by Northern analysis of multiple adult rat tissues (Higashide *et al.* 1996) therefore cannot be precluded. This has even been indirectly acknowledged by the group who performed the original screen, because their genomic analysis reveals that the UNC119 promoter consists of several GC boxes instead of the classical TATA box (Higashide and Inana 1999). This structure is typical of non-abundant, widely expressed housekeeping genes (Mitchell and Tjian 1989). The existence of lower level temporal patterns of UNC119 expression could therefore potentially explain the functional conservation of UNC-119 proteins despite apparently vast differences in localization and suggest non-retinal roles for UNC119 in mammals during development.

Another possibility that could explain the contrasting patterns of expression between the unc-119 family members would be the existence of other mammalian unc-119 homologues. Many systematic gene duplication events have taken place during the adaptive radiation of vertebrates from their ancestors, and they have been critical to development of the vertebrate lineage (Cooke *et al.* 1997).

A variety of fates are possible for duplicated genes. Due to redundant functions, most are lost over time, because they accumulate degenerating mutations unless there is some sort of selective pressure that maintains them (Force *et al.* 1998). The acquisition of

novel expression sites for developmental genes appears to be the major mechanism of duplicate gene retention within complex metazoans, because this provides an opportunity for new gene roles that allow the progressive extension of development (Cooke *et al.* 1997). However, partitioning of ancestral functions (sub-functionalization) due to complementary degenerative mutations following duplication is also common. In such cases, expression of each duplicate occurs in different subsets of the temporal and spatial expression pattern of the ancestral gene, and the sum of their expression patterns mirrors the expression pattern of the original gene (Force *et al.* 1998).

The vast differences in unc-119 expression and protein localization observed in different organisms to date could therefore potentially be explained by the presence of at least one other mammalian homologue whose expression and localization complements that observed for UNC119. Furthermore, if additional homologues do exist, a comparison of their properties could reveal the type of selective pressure maintaining each one and lead to insights into the timing and sequential progression of development.

Although different unc-119 homologues with distinct properties have been identified in zebrafish, their existence in fish does not necessarily suggest that multiple homologues exist in mammals. Many duplicate genes are unique to ray-finned fish, because this branch of vertebrates has undergone a genome duplication event since their divergence from tetrapods, the lineage to which mammals belong (Postlethwait *et al.* 1998).

Even though there has been less time for the evolution of changes in duplicate genes since the more recent fish duplication event, there is precedence for specific branching of function between newly duplicated genes in zebrafish. Many examples of specific sub-functionalization in zebrafish exist, such as the divergence of the *engrailed (eng)* genes (Force *et al.* 1998). In tetrapods (mouse and chicken), *En1* is expressed in the developing pectoral appendage bud and in specific neurons of the hindbrain and spine (Joyner and Martin 1987; Davis *et al.* 1991; Gardner and Barald 1992). In fish, however, *eng1* is found in the appendage bud while *eng1b* is expressed in the neurons (Force *et al.* 1998). The existence of other unc-119 homologues in mammals is therefore not predictable based on the existence of *Danio rerio unc119* and *unc119b*, because it is

unknown whether the presumed unc-119 duplication event preceeded or followed the ray-finned fish genome duplication.

1.8 Project objectives and approaches taken

The objectives of this project are to determine the cell biological function of UNC119 and to delineate the evolution of the unc-119 family. By mapping sites of activity and interaction of UNC-119 and other possible homologues in mammals, and by assessing its sub-cellular localization in cultured expressing cells, this branch of the highly conserved family will be further characterized. The information obtained will then be integrated with the data already available from other organisms to elucidate the cellular function(s) and mechanisms of activity of UNC-119 proteins.

Mus musculus (mouse), *Rattus norvegicus* (rat) and *Homo sapiens* (human) have been selected as mammals for study. These closely related genomes have been sequenced, and their genomic, proteomic, and syntenic data are publicly available, making genomic and comparative analyses possible. Mouse is a model mammalian organism for experiments requiring intact embryonic and adult tissue, because it is widely available, and on-site animal care facilities are in place and accessible. Furthermore, the feasibility of this project is greatest using this system, because successful parallel types of studies using mouse embryos of various stages are frequently reported in current literature.

To pinpoint the sub-cellular distribution of UNC119, cultured cells are ideal, because they grow in single layers against labware, and they can be easily immunologically stained and visualized using high power fluorescent optics. The rat PC-12 cell line is readily available and has been widely studied. This line was chosen from an extensive suite of available samples, because these pheochromocytoma derived cells are reversibly inducible to differentiate into neurons and project axons *in vitro*. Provided that they express UNC119, the use of such reversible cells would be most informative for this study because protein localization and distribution could be followed throughout their differentiation, growth, and development.

To attain these objectives, the 3 areas this project has focused on are:

1. Identification and genomic analysis of novel mammalian UNC-119 family members.

2. Characterization of non-retinal mammalian expression and protein localization of UNC119 and novel homologue(s).

3. Identification of proteins physically interacting with UNC119.

Materials and Methods

2.1 Northern analysis

Commercial multi-tissue fetal and adult human Northern blots (Human Fetal MTN® Blot II, Human 12-Lane Multiple Tissue Northern (MTN™) Blot by Clontech, Inc.) were probed using radiolabeled single-strand RNAs prepared as described below. A β-actin probe (supplied with the blot) was used as a loading control.

Probes were synthesized using the T7 Quickprime kit (Pharmacia Biotech) according to standard procedures using labeled α-^{32}P dCTP. The UNC119 probe used was a 688 base BamHI fragment of pDP#MM092R (Maduro *et al.* 2000). The UNC119B probe used was the ~600 base NotI/EcoRI insert from IMAGE clone 259613 (gi:1165809). Blots were hybridized overnight at 65°C in Hybrizol II (Serologicals Corporation) and washed at the same temperature. They were washed twice for 15 minutes with 2X SSC followed by a 30 minute wash with 2X SSC/0.1% SDS and a 10 minute wash with 0.1X SSC/0.1% SDS. Stripping of the blot was performed as indicated in the manual supplied with the blots.

2.2 Mouse embryo collection and storage

For *in situ* hybridization experiments, embryos were dissected from pregnant CD-1® (ICR) BR mice (obtained from Charles River, Inc.) and immediately fixed in DEPC treated 4% paraformaldehyde at 4°C for different lengths of time depending on embryo size. Fixation times were 26 hours for embryonic day 10 (E10.5) embryos, 50 hours for E11.5, 73 hours for E12.5, and 69 hours for E13.5. They were then washed 3 times for 30 minutes in DEPC treated PBS at room temperature and dehydrated in a series of 30%, 50%, and 70% ethanol in DEPC treated PBS at 4°C for an hour per step. Embryos were stored in 70% ethanol-DEPC-PBS at 4°C.

For the immuno-chemistry of the non-commercial sections presented in this thesis, E11.5 embryos were dissected and fixed overnight as described above. Fixative was injected directly into the heads and livers of the embryos 1 hour post-collection using

a 27½ gauge needle. Embryos were decalcified for an hour in RDO (Du Page Kinetic Laboratories, Inc.) at room temperature, washed 3 times for 30 minutes each in DEPC treated PBS (shaking), and dehydrated in a series of sucrose solutions. Each was bathed in 12% sucrose at room temperature for 6 hours and transferred to 18% sucrose overnight. Embryos were rapidly frozen in Tissue-Tek® embedding medium (Sakura Finetek USA, Inc.) over a mixture of isopentane and liquid nitrogen and immediately stored at –80°C. Cryostat sectioning was performed to a thickness of 12μm and mounted on poly-L lysine coated slides (Erie Scientific Company, Inc.). Only the most para-sagittal sections were intact and used.

2.3 Whole mount *in situ* hybridization

Staining of whole embryos was performed using a digoxigenin based protocol described in Appendix 1. In brief, embryos were rehydrated in a series of ethanol gradients, rinsed twice in DEPC treated PBT, and bleached in 6% H_2O_2. They were then rinsed 3 times in DEPC treated PBT, permeabilized in 20μg/ml proteinase K, rinsed 3 times, fixed in 4% paraformaldehyde / 0.2% glutaraldehyde, and rinsed 4 more times. Embryos were then bathed twice for at least an hour in hybridization solution (50% formamide, 5X SSC pH4.5, 1% SDS, 50μg/ml yeast RNA, 50μg/ml heparin) at 68°C prior to hybridization overnight at the same temperature.

Following hybridization, embryos were washed twice for 30 minutes at 68°C (in 50% formamide, 5XSSC pH4.5, 1% SDS), equilibrated, and RNase A (100μg/ml) treated twice in RNase buffer (0.5M NaCl, 10mM Tris-Cl pH7.5, 0.1% Tween 20) for 30 minutes at 37°C. Following rinsing, embryos were washed 2 more times at 68°C (in 50% formamide, 2XSSC (pH4.5)), equilibrated in TBT, blocked for at least 2.5 hours in 10% heat inactivated mouse serum (diluted in TBT), and incubated at 4°C overnight in a 1:2000 dilution of pre-absorbed α-DIG antibody (in TBT, 1% mouse serum, 1% blocking reagent and 2mM levamisole). Several prolonged washes in TBT supplemented with 2mM levamisole were then performed to remove unbound antibody.

Following equilibration in developing AP solution (100mM Tris-Cl pH9.5, 100mM NaCl, 50mM $MgCl_2$, 2mM levamisole, 0.1% Tween 20), colour development

was performed in the dark using 5µl NBT and 3.75µl BCIP (Promega, Inc.) per ml of AP solution. Once colour was optimal, the reaction was stopped with 2 quick and several prolonged washes using 20mM EDTA in PBT. Finally, embryos were fixed for an hour in 4% paraformaldehyde in PBT, rinsed, and cleared using a methanol dehydration gradient series (25%, 50%, 75%, 100% twice). Embryos were then rehydrated, rinsed, and stored in PBT at 4°C.

Probes were synthesized using the Megascript RNA polymerase T7 and SP6 kits (Ambion, Inc.) and the constructs described in Appendix 3. For UNC119 probes, plasmid was linearized using SalI and NotI, and labeled using the T7 and SP6 RNA polymerases for antisense and sense probes respectively. For UNC119B, probes were prepared using SalI restriction enzyme and T7 polymerase for antisense, and XbaI and SP6 for the sense probe. *Pax1* antisense probes were prepared using T7 polymerase and a NotI linearized template.

2.4 *In situ* hybridization of thin tissue sections

Staining of 7µm E8.5, E10.5, and E14.5 murine sections (Novagen, Inc.) was performed using the non-radioactive protocol detailed in Appendix 2. In brief, slides were dewaxed in toluene, rehydrated in a series of ethanol gradients, rinsed once in DEPC treated water and then 2 more times in DEPC treated PBS. Sections were then fixed in 4% formaldehyde for 10 minutes and treated twice to remove RNases for 15 minutes each in PBS containing 0.1% active DEPC. They were permeabilized in 20µg/ml proteinase K diluted in TE at 37°C for 20 seconds and immediately equilibrated twice in DEPC treated 5X SSC. They were bathed in hybridization solution (50% formamide, 5X SSC, 500µg/ml salmon sperm DNA) for at least 2 hours at 60°C and 50°C (for UNC119 and UNC119B respectively) prior to hybridization overnight at the same temperature. The probes used were identical to those used in whole mount *in situ* hybridizations.

Following hybridization, sections were washed for 30 minutes in 2X SSC at room temperature and then for 60 minutes at hybridization temperature. They were washed a third time in 0.1X SSC for 60 minutes at hybridization temperature. They were then equilibrated twice in buffer (10mM Tris pH7.5, 150mM NaCl) and incubated in a 1:2000

dilution of α-DIG antibody (in the same buffer supplemented with 1% blocking reagent (Roche, Inc.)). Following 2 washes in the same buffer, they were equilibrated twice in buffer 2 (100mM Tris pH9.5, 100mM NaCl, 50mM $MgCl_2$) and developed using NBT/BCIP (Promega, Inc.) at 4.5µl NBT and 3.5µl BCIP per ml of buffer 2. Once colour was optimal, reactions were stopped in TE pH8.0, sections were counterstained in methyl green, and dehydrated in a series of ethanol washes of increasing concentrations (35%, 50%, 75%, 95% and 100%). They were finally washed in toluene and mounted in DPX media (BDH laboratory supplies, Inc.).

2.5 Antibody production

For UNC119, the glutathione S-transferase (GST)-tagged full-length human UNC119 expression construct ufp-5 (Maduro 1998) was expressed in *E. coli* strain BL21(DE3) under standard conditions and purified against Glutathione Sepharose 4B (Pharmacia Biotech) according to manufacturer's recommendations (GST Fusion System handbook 1993). Rabbit antiserum to the fusion protein was raised by immunization of 2 rabbits (OF3 and OF5) using methods described elsewhere (Huet 1998).

A human UNC119B expression construct containing the carboxyl 243 amino acid residues fused in frame to GST was created by ligating an oligonucleotide adapter (a dimer of 5'-GGCCGCGGATCCGC-3') into NotI linerarized IMAGE clone 259613 5' (gi:1165809). The 259613 5' insert/adapter was isolated using NotI and EcoRI and ligated into pGEX-2T (Pharmacia Biotech). The cloning site was sequenced to confirm that UNC119B was ligated in frame using the 5'GST vector sequencing primer 5'-GGGCTGGCAAGCCACGTTTGGTG-3'(Pharmacia Biotech). This construct was then transformed into BL21(DE3), expressed, and purified as described above for ufp-5. Antiserum to the fusion protein was raised by immunization of 2 rabbits (11-5 and 11-6) as described elsewhere (Huet 1998).

2.6 Antiserum purification

Large quantities of purified GST-tagged fusion proteins (100-2000μg) were size fractionated and transferred to membranes by Western blotting. The portions of the membranes containing the largest, brightest bands for each protein (indicated by the arrows in figures 21A and 21B) were blocked overnight at 4°C in 3% BSA diluted in PBS. Following a quick rinse in PBS, strips were incubated overnight at 4°C with aliquots (~2ml) of un-purified serum diluted to a final concentration of 20% in PBS. Strips were washed 3 times for 10 minutes in PBS, and then bound epitopes were eluted from the membranes by incubation for 2 minutes in 2ml acidic glycine (0.1M glycine-HCl pH2.3) and immediately neutralized by addition of 0.5ml neutralization buffer (0.09M Tris pH8.8, 0.1M NaCl). The eluted fractions were washed 3 times at ~2000rpm in a concentration column using 10-15ml PBS, and purified epitopes were concentrated to a final volume of ~0.3ml and stored at -80°C.

2.7 Western analysis

Western blotting and analyses were perfomed according to standard procedures using nitrocelluse membranes and detected by enchanced chemiluninescence using the Supersignal Kit (Pierce, Inc.) and film exposure.

2.8 Mammalian tissue collection for protein extraction

Bovine brain tissue and intact eyes were collected from a local slaughterhouse (Alsask Food Processors, Inc.) on ice, retinas were dissected on ice, and proteins were extracted from the tissues as described elsewhere (Higashide 1998).

Whole mouse embryos were dissected from pregnant CD-1® (ICR) BR mice (obtained from Charles River, Inc.), placed on ice, and immediately extracted and homogenized in the same buffer used for the bovine extracts that was freshly supplemented with a complete protease inhibitor cocktail pellet (Boehringer Manheim, Inc.) immediately before use.

2.9 Non-denaturing protein extraction

Whole E14.5 mouse embryos were homogenized on ice in 3-5 volumes of fresh extraction buffer (25mM Tris-HCl pH7.5, 100mM NaCl, 0.5% CHAPS, 50mM NaF, 1mM EDTA, 1mM EGTA, 10μg/ml leupeptin, 10μg/ml antipain, 0.5mM TLCK, 0.7μg/ml chymostatin, 10μg/ml benzamidine, 1mM PMSF) and pelleted at 4°C by centrifugation. Cleared lysates were immediately used or stored at –20°C.

2.10 Immuno-chemistry

For staining using frozen tissue, slides were thawed at room temperature, dried for an hour at 40°C, fixed in acetone for 10 minutes at –20°C and briefly air dried. Staining was performed as described in detail in Appendix 4. In brief, sections were washed, and endogenous peroxidases in the tissues were blocked in a 3% hydrogen peroxide solution diluted in methanol. They were then washed and blocked (6% goat serum, 2.4% protease-free BSA) with Avidin (4 drops/ml). Sections were incubated overnight at 4°C with α-UNC119 OF3 or OF5 polyclonal antiserum at various dilutions (1:50-1:500). They were then washed, incubated in a 1:200 dilution of α-rabbit IgG (raised in goat) for an hour, washed, bound to a biotinylated-peroxidase enzyme for signal amplification, washed, and developed using a DAB substrate/nickel mixture.

Prior to immuno-chemical staining of paraffin embedded tissue sections, slides were dewaxed in 2 changes of xylene for 5 minutes each and rehydrated in a series of ethanol changes (100% twice, 90%, 70% and 50%) for 2 minutes each. Following a rinse in distilled water, antigen unmasking was performed using a 0.1M citrate buffer (made using anhydrous citric acid and adjusted to a pH of 6.0 using sodium hydroxide). Slides were immersed in buffer, brought to a boil and microwaved for 5 additional minutes, and cooled to room temperature. Slides were then removed from buffer and immediately stained as described for the frozen sections with the exception that all washes were performed using PBT. Blocking was performed overnight at 4°C and primary antiserum incubations were performed for 2.5 hours at room temperature.

For the α-UNC119 treated sections, the purified OF3 antiserum and pre-absorbed controls (0.1μg UNC119-GST per μl antiserum) were diluted and incubated overnight at 4°C prior to application. Biotin was added to the mixtures immediately before use.

2.11 Cell culture

PC-12 cells were purchased from American Type Culture Collection (ATCC) (catalogue # CRL-1721). Aliquots of PC-12 cells were revived and cultured as recommended by the manufacturer in Kaighn's modification of Ham's F12 medium (F12K) supplemented with 15% horse serum, 2.5% fetal bovine serum (ATCC) and 1% L-glutamine. After splitting, several vials were preserved at –80°C for future study.

Total protein was extracted from a 50ml cell suspension by resuspension of pelleted cells in 500μl 2X SDS sample buffer (Ausubel *et al.* 1987) and 100μl PBS. The sample was boiled and stored at –20°C.

2.12 Tagged-protein pull-down assay

Approximately 1mg of purified UNC119-GST protein was incubated with gentle agitation for 30 minutes at room temperature in a 1.5ml slurry of prepared Glutathione Sepharose 4B (Amersham Biosciences, Inc.). The UNC119-GST bound beads were then pelleted (500 X g for 5 minutes), and washed 3 times with PBS and 2 more times with the extraction buffer used to obtain the non-denatured murine protein lysates. Approximately 1mg of non-denatured protein extract was pre-cleared by agitation with 500μl of unbound glutathione slurry for 45 minutes at 4°C to remove any non-specific binding proteins. The pre-cleared murine protein lysate was then incubated with the UNC119-GST bound beads for 45 minutes at room temperature with agitation, pelleted, and washed 5 times with 5-6ml extraction buffer. Bound proteins were eluted in 2X SDS sample buffer (Ausubel *et al.* 1987) and resolved on an SDS-PAGE protein gel.

2.13 Antibody immuno-precipitation assay

30μl of protein A-Sepharose 4 fast flow beads (Amersham Biosciences, Inc.) were pelleted, washed 3 times with 10X volume of non-denaturing protein extraction buffer, and resuspended in fresh buffer to make a 50% slurry. 8μl of each α-UNC119 antiserum (OF3 and OF5) was diluted in 500μl extraction buffer and incubated by gentle agitation with 20μl of bead slurry for an hour at 4°C. Conjugated beads were washed 3 times in 1ml extraction buffer and resuspended to form a 50% slurry. The coated beads were then added to ~5mg (227μl) of non-denatured murine protein lysate and 500μl extraction buffer. This mixture was rotated overnight at 4°C, pelleted, and washed 4 times in 1ml extraction buffer at 4°C to remove any unbound proteins from the mixture. Residual extraction buffer was aspirated using a 27½ gauge needle and any remaining proteins were eluted into 30μl 2X SDS sample buffer (Ausubel *et al.* 1987).

Results

3.1 UNC119B: A second UNC119 homologue in mammals

Prior to the initiation of this project, an expressed sequence tag (EST) (gi:1165809) had been located from gene expression databases that showed similarity to human UNC119. ESTs are short DNA sequence reads (typically 300-500 base pairs) obtained from cDNA that represent a snapshot of genes expressed in a given cell type or at a given developmental stage. EST 1165809 is 71% identical to a region of the UNC119 transcript (gi:4103879) as demonstrated in figure 9A. Despite the extensive similarity between the sequences, differences between the EST and the known UNC119 cDNA are significant enough to suggest the presence of a second UNC119 homologue in the human genome. This putative family member is also expressed outside of the retina, because the EST was sequenced from placental tissue.

Further searches against public EST databases reveal a large group of overlapping sequences corresponding to this predicted second UNC119-like gene. The ESTs are derived from a variety of tissues and cell lines, including fetal heart and brain, fibroblasts, and tumors derived from ovary, colon, lung and adrenal gland. This transcript is therefore widely expressed. An assembly of the overlapping ESTs reveals that their combined sequence contains a large open reading frame that, when translated, encodes a protein 56% identical to human UNC119, as shown in figure 9B. This sequence has been submitted to the HUGO Nomenclature Committee and the gene has been officially named "unc119 (*C. elegans*) homolog B". It is represented by the symbol UNC119B.

Unfortunately, even though multiple ESTs corresponding to the putative gene exist, none identified at this time contain the 5' end of the open reading frame or the 5' untranslated region of the UNC119B transcript. The open reading frame of UNC119B extends to the 5' most end of the sequence assembly, and it does not contain the conserved translational start site 'ATG^{start}' motif or any upstream transcriptional motifs.

During the course of this project, database searches have been regularly conducted to identify new UNC119B mRNA transcripts and search for its 5' end. Periodic searches have been necessary, because several clones initially identified were

subsequently retracted from the database, either at the submitter's request or by standard genome annotation processing. One search identified a predicted UNC119B mRNA (gi:22057100) that was based on computational analysis of the human genome. When this predicted sequence was used to search for authentic clones, a partial transcript derived from neuroepithelioma tissue (gi:16552587) was identified, and it was mapped against the human genome to determine the gene structure of UNC119B and search for its 5'end. It was localized to chromosome 12q24.31 with a 5 exon structure. Examination of upstream genomic sequence using various prediction software programs, however, did not reveal any probable in frame 'ATG^{start}' codons (Kozak 1996) for UNC119B.

Recent updates to the human genome chromosome 12 sequence indicate that the mapped transcript actually corresponds to MGC5139 (gi:29744858), a small predicted protein located immediately adjacent to UNC119B and translated in the opposite orientation. Indeed, even though the partial neuroepithelioma transcript (gi:16552587) and the UNC119B DNA sequence assembled from overlapping ESTs both matched the predicted UNC119B mRNA (gi:22057100) initially identified, their sequences do not overlap, and the predicted mRNA has been retracted from the database. The retracted 4 Kb predicted transcript erroneously spanned both loci.

Most recent searches (April 2003) map the assembled UNC119B sequence to a genomic chromosome 12 BAC clone (gi:22538699) and to the human chromosome 12q24.31 supercontig (gi:29744945) shown in figure 10A. An alignment of the assembled coding sequence to both genomic clones suggests a 5 exon structure for UNC119B, depicted in figure 10B. This structure is consistent with the conserved genomic structure of other unc-119 family members, because the *Caenorhabditis* and *Drosophila* unc-119 genes, as well as human UNC119 and zebrafish Unc119 are all structurally organized into 5 exons (Maduro and Pilgrim 1996; Maduro *et al.* 2000; Higashide and Inana 1999).

The DNA sequences immediately 5' of the UNC119B coding sequence in both genomic chromosome 12 clones also contain an in-frame potential translation initiation 'ATG^{start}' codon which would predict the existence of 8 extra amino acids (shown in figure 11A) and a total size of 251 amino acids for the UNC119B protein. Although this putative initiation codon is surrounded by the residues "GGAGCGatgA" which differ in

several positions from the sequence motif "GCCACCatgG" typical for vertebate transcripts (Kozak 1996), evidence from other mammalian homologues suggest that this is the authentic site of translation initiation.

Similar to human UNC119, UNC119B is very highly conserved in mammals, as shown in figure 11B. Early database searches identified a partial mouse mammary tumor mRNA transcript (gi:14250370). This clone was ordered and subsequently used for the UNC119B expression studies described later. Very recently (April 2003), several other cDNAs corresponding to the mouse transcript have become available, and these include a full-length transcript (gi:26349818) derived from a neonate cerebellum library. Another partially overlapping transcript obtained from embryonic parthenogenote tissue (gi:26337100) has also been identified. A predicted full-length transcript (gi:27666035) corresponding to rat UNC119B has also become available, though its authenticity remains to be determined. The translated open reading frame of each rodent protein is 90% identical to human UNC119B, and they are 97% identical to each other, confirming the extensive conservation of this newly identified mammalian homologue.

Mouse UNC119B maps to mouse chromosome 5E5, and, as demonstrated in figure 12, also appears to have a 5 exon structure based on the full-length clone. Only 4 regions of identity of mouse UNC119B map to chromosome 5E5, however a 173 nucleotide region (positions 538-711 of gi:26349818 or 1-178 of gi:14250370) representing putative exon 4 does not correspond to any portion of the most recent version (April 18, 2003) of the mouse genome. The un-mapped sequence represents a portion of the UNC119B open reading frame, suggesting that this version of the mouse genome either contains a small gap at this location or has yet to be properly assembled. Future searches as the genome annotation becomes more complete should resolve this discrepancy. A region of 1353 bases at the distal end of the 3' untranslated region of the UNC119B transcript also matches mouse chromosome 18, suggesting that a repeat sequence is present at this location.

Human to murine syntenic searches have also been conducted and indicate that UNC119B is located in a region of conserved synteny between the human and mouse chromosomes based on the Jackson Laboratory Mouse Genome Informatics database (http://www.informatics.jax.org/searches/homology_report.cgi). A mammalian orthology

comparative query between human chromosome 12 and mouse chromosome 5 demonstrates that the region of conserved synteny between these chromosomes is extensive, spanning at least 86 different loci on either side of UNC119B. Furthermore, tightly linked genes on either side of human UNC119B have been identified that have mouse orthologues mapping to the corresponding regions surrounding UNC119B on the mouse chromosome. As demonstrated in figure 10A, the human gene acyl-Coenzyme A dehydrogenase C-2 to C-3 short chain (ACADS) (gi:337927) is just distal of human UNC119B at 12q22.12-qter, and the human calcium binding protein 1 (CABP1) gene (gi:6708059) is located immediately proximal to UNC119B at 12q24.31. The mouse orthologues for both genes map to regions adjacent to mouse UNC119B as shown in figure 12. Acads (gi:6680619) and Cabp1 (gi:7304938) map to mouse chromosome 5 at 65.0 cM and to cytoband F, respectively.

The existence of multiple murine UNC119B transcripts suggests that alternate forms of UNC119B may exist. As demonstrated in figure 13, a cDNA clone from 9.5 day embryo parthenogenote tissue (gi:26337100) has also been identified. This clone shares the first and second exons with the full-length transcript (gi:26349818), however its second exon extends 957 bases further along the chromosome 5 genomic sequence than does the full-length transcript, into the intronic region between exons 2 and 3 of UNC119B. The first 114 amino acids predicted by this clone's open reading frame are consistent with those identified in the other UNC119B transcripts, however its open reading frame continues into the unique portion of the clone, predicting a protein of 237 amino acids in length. This unique sequence shares no homology to any other identified sequences in mouse or in other organisms, therefore its authenticity remains questionable. Another short 5' UNC119B clone predicted by computational analysis (gi:28547940) also supports the location of the first intron. Because this clone appears to contain sequence gaps, however, it has been disregarded during further analysis.

Sequence comparisons of both mammalian UNC-119 homologues to the zebrafish UNC-119 proteins have also been conducted. Although it might be expected that each mammalian homologue would correspond to its own respective zebrafish orthologue, both UNC119 and UNC119B are more similar to Unc119b (70% and 66% identical respectively) than to Unc119 (66% and 54% identical respectively). However, despite

higher similarity to Unc119b, regions of conserved synteny between human and fish chromosomes indicate that UNC119 corresponds to Unc119. Syntenic analysis of UNC119B has not been successful to date but should become possible in the near future as the fish database becomes more complete. A third potential unc-119 homologue in zebrafish was also identified by Southern analysis. Partial sequencing demonstrates that it is related to the unc-119 family, however it is unclear at this time whether or not it is expressed (Manning, personal communication).

To better compare the number of existing UNC-119 family members in vertebrates, the largely sequenced Fugu (a puffer-fish) genome has also been consulted. Based on un-assembled genomic data, it appears as though there are at least 3 distinct regions in the Fugu genome that potentially represent UNC-119 homologues based on sequence similarity to the human UNC119 transcript. A putative translation of each Fugu sequence reveals that 2 are more similar to UNC119B and that the other is more similar to UNC119. Out of the 107 amino acid differences between human UNC119 and UNC119B, the first Fugu region matches specifically to one of the human UNC-119 proteins 48% of the time with 78% of these matches corresponding to UNC119B (shown in Table 1). For the second region, the predicted Fugu protein specifically matches one human homologue 53% of the time, with 83% of the matches corresponding to UNC119B. In contrast, the third putative homologue matches one of the unconserved residues between the 2 human UNC-119 proteins 38% of the time, with 60% of the matches to corresponding to UNC119. As the Fugu genome annotation becomes more complete and expression data becomes available, it will be interesting to determine if Fugu has 2 authentic UNC119B-like homologues and 1 UNC119-like homologue.

Domain analysis of the predicted UNC119B protein has also been conducted to further characterize this product. Based on its amino acid sequence, UNC119B has been assigned a probability of 74% that its sub-cellular localization is cytoplasmic by PROFsec software (Emanuelsson *et al.* 2000). It appears to adopt a non-globular conformation (Rost, personal communication) with a core to surface ratio of 35 to 65 based on a GLOBE prediction (Rost and Sander 1993; Rost *et al.* 1996). PROFsec software also predicts that its secondary structure consists of interspersed alpha helices (30%) throughout the protein with some beta sheets (22%) in the carboxyl portion (Rost

and Sander 1993; Rost *et al.* 1996). A similar distribution (34% or 36% alpha helices and 14% or 18% beta sheets) is also predicted using independent (GOR4 or HNN) software (Combet *et al.* 2000), providing strong evidence that UNC119B is a protein that adopts a mixed conformation. No transmembrane domain is predicted. Of the cysteine residues present in the protein, the only cysteine predicted (with a probability of 0.62) by CYSPRED (Fariselli *et al.* 1999) to form disulfide bridges is located at position 85. This residue is conserved in all members of the UNC-119 family.

In contrast to several N-terminal predicted O-glycosylation sites of serine or threonine residues predicted for UNC119 by NetOGlyc 2.0, YinOYang 1.2, only a few very weak potential sites are predicted for UNC119B. The strongest predicted glycosylation sites for UNC119 are serines at positions 16 and 35, with several weaker sites predicted at positions 21, 24, 37, 39 and 41 (Hansen *et al.* 1995, 1997, 1998; Gupta *et al.* 2003). None of these residues are conserved with UNC119B, as demonstrated in figure 9. However, weak N-myristoylation signals are predicted in the N-terminal of UNC119B by some prediction software (ProSite) (Hofmann *et al.* 1999), and these are also in non-conserved regions of the proteins. PESTfind analysis also suggests that both proteins contain possible PEST proteolytic signals that could target the proteins for rapid destruction (Rogers *et al.* 1986; Rechsteiner and Rogers 1996).

One type of reversible post-translational modification that is critical to regulation of activity or of the structural conformation of many proteins is phosphorylation of key residues. Several potential phosphorylation sites of serine, threonine, and tyrosine residues are predicted for both UNC119 and UNC119B. Of the UNC119B residues (with the predicted 5' end inclusive) predicted by multiple software (NetPhos 2.0, ProSite) to most likely be phosphorylated (serine 119, threonine 67, tyrosines at positions 205, 223 and to a lesser extent 177 and 245) (Blom *et al.* 1999; Hofmann *et al.* 1999), all of the tyrosine residues are conserved in UNC-119 at least throughout vertebrates, and the serine and threonine residues are specific to all vertebrate UNC-119 "B" proteins.

3.2 UNC119 and UNC119B are widely expressed during development

Previous reports have clearly demonstrated strong retinal expression of UNC119 in adult mammalian tissue (Higashide *et al.* 1998) with suggestion of very faint expression elsewhere after prolonged exposures of Northern blots (Swanson *et al.* 1998). Because no solid evidence of non-retinal UNC119 expression existed prior to the initiation of this project, and the only known information regarding UNC119B expression was a single corresponding EST sequenced from a human placental tissue library, Northern analysis was performed to test for non-retinal expression of both genes and to obtain a general overview of their temporal and spatial expression. As demonstrated in figure 14, both UNC119 and UNC119B are widely expressed in human fetal tissues. Consistent with the previously reported transcript size (Higashide *et al.* 1998), strong hybridization of UNC119 to a 1.37 Kb mRNA was observed in the 4 fetal tissues at a relatively similar level (figure 14A). Fainter ~2 Kb and 7.5 Kb bands are also present, and these could represent alternative splice variants or be artifacts of cross-hybridization to other transcripts. The fainter 2 Kb band is consistent with previous observations and has been characterized as an incompletely processed transcript by others (Swanson *et al.* 1998). The nature of the larger transcript is unknown.

As shown in figure 14B, UNC119B also hybridizes to a transcript in every fetal tissue tested at least at trace levels. This much larger transcript is approximately 4.5 Kb, smaller than the faint 7.5 Kb transcript detected by the UNC119 probe. Extremely faint cross-reaction of the UNC119 probe with UNC119B is possible, because a band consistent with the size of the UNC119B transcript is visible below the 7.5 Kb band. Although the overall expression levels of UNC119B appear lower than UNC119, differences in specific probe radioactivities make it difficult to quantify this difference.

A multi-tissue human adult Northern blot has also been probed with each transcript (data not shown), and no expression was detected in any of the 12 (non-retinal) tissues tested. Probing of the blot with multiple control transcripts following the UNC119B experiments was also performed to rule out the possibility of membrane deterioration. However, the possibility of low level adult expression has not been excluded, because control probes are designed to yield a strong signal against an

abundant transcript, therefore they could still be detectable on a repeatedly used and stripped blot on which rare transcript detection would not be visible.

Although Northern analysis is an excellent tool for obtaining a rapid overview of tissues expressing a transcript of interest, a technique such as *in situ* hybridization (ISH) is required to further define expression patterns. Unlike Northern blots, ISH distinguishes differences in expression between the various cell types or regions within a tissue, and because the probes are hybridized to an individual sample derived from a single organism, temporal changes in expression may also be determined. In contrast, Northern blots are prepared using pooled mRNA. The fetal blots are also fabricated using unstaged tissues and therefore are not informative in determining developmental temporal expression patterns.

To characterize the expression patterns of UNC119 and UNC119B throughout development, ISH was performed on both whole mount and sagittal sections of developing mouse embryos. Several stages of development have been examined.

3.3 Whole mount *in situ* hybridization shows ubiquitous expression

To obtain a temporal and spatial measure of expression, mouse embryos representative of various stages of development (embryonic days 10-15) were collected and stained for UNC119, UNC119B, and unrelated controls. A "no probe" control was also performed and did not reveal any background staining.

At embryonic day 10 (E10.5), a stage characterized by rapid growth and tissue differentiation, there is essentially ubiquitous expression (demonstrated in figure 15) of both UNC119 and UNC119B. It appears as though there may be a slight up-regulation of both genes in the limb buds and pharyngeal arches and exclusion (most evident for UNC119B) in the heart. However, it remains possible that this observation is an artifact of differences in tissue shape and embryo fragility. Based on this data, it is difficult to determine if there are differences between the expression patterns of UNC119 and UNC119B.

As development continues, it appears as though both transcripts become progressively localized. As shown in figure 16, ubiquitous expression of UNC119 and

UNC119B continues to be apparent by E11.5. Antisense stained embryos are much more darkly stained overall compared to those stained with sense probes or with antisense probes of unrelated transcripts. Embryos hybridized using a *Pax1* probe (which specifically stains somites (Deutsch *et al.* 1988)) do not have the pink hue apparent throughout the UNC119 or UNC119B stained embryos, demonstrating that this overall coloration does not correspond to non-specific background staining. Similar observations using another transcript, *IPW*, (data not shown) confirm these results.

By E12.5, however, expression of both homologues appears to become more localized. As shown in figure 17, UNC119 expression appears concentrated in the developing limbs, facial area, and base of the head/cervical region with faint expression along the spinal region when compared to the sense control. UNC119B is expressed in similar regions, although it is distributed more prominently in the posterior regions of the body with little or no staining of the facial structures. However, because a restriction in staining distribution and reduction in signal intensity is also observed for *Pax1* stained embryos, this signal localization is likely an artifact of the thickened tissue characteristic of older embryos. Staining in brain tissue is also observed in all embryos and is likely a consequence of probe trapping in the large cavities present at this stage.

This becomes more evident as development proceeds further, because staining of E13.5-E15.5 embryos yields inconsistent and variable patterns that are restricted to the superficial tissues of the body regardless of the probe applied (data not shown). A review of current literature also suggests that successful staining of older embryos is typically performed on thin tissue sections obtained from representative regions of the specimen, therefore this approach has been adopted to obtain a more complete analysis of UNC119 and UNC119B expression during development.

3.4 *In situ* hybridization of mouse embryo sagittal sections

To further characterize the cell and tissue types expressing UNC119 and UNC119B throughout development, ISH has been performed on randomly oriented sections of embryonic day 8 (E8.5) and 10 (E10.5) mouse embryos *in utero*, and on para-sagittal sections of E14.5 embryos. The mouse gestation period is 19-21 days.

In younger embryos (E8.5 and E10.5), little expression of either probe above background is observed. However, as shown in figures 18A, 18E, 19A, and 19C, prominent staining of certain extra-embryonic tissues is consistently observed in all sets of sections examined to date (at least 4 for each probe). In the trophoblast, some dividing cells closest to the embryo are strongly stained, though a progressive reduction in staining of these cells is observed as their infiltration into the uterine mucosa increases. Intermittent staining of several larger cells is also observed. Without further staining analysis, it is unclear whether these large cells are giant cells of the trophoblast or decidual cells that are undergoing changes following the decidual cell reaction. These cells, located in the uterine tissue of pregnant mothers, undergo several changes during and following implantation of an embryo (Kaufman and Bard 1999). Examination of stained tissue at high magnification (figures 18C and 18D) reveals that the staining is very diffuse, cellular, and non-uniform. Although the presence of extra-embryonic staining would be expected based on the presence of UNC119B ESTs sequenced from placental tissue in the public database, the significance of the staining is highly questionable given that similar strong staining patterns are also observed on the sense probe controls.

Initial experiments suggested that several embryonic tissues of para-sagittal E14.5 sections stain for both UNC119 and UNC119B. As demonstrated in figures 18G and 19E, the most prominent sites of detection are in the ossifying bones throughout the embryo. Darkly stained ossification centers include the shoulder girdle, the spinal vertebral bodies, the first rib, the mandibular and maxillar bones of the face, and several others dispersed throughout the embryo. The cells expressing the UNC119 homologues appear to be those actively undergoing differentiation during the ossification process, because the stained structures correspond to those that ossify during the 14^{th} and 15^{th} days of gestation (Rugh 1968). Structures that are still cartilaginous at this stage do not show evidence of expression. Further evidence that the stained cells are actively ossifying is their predominantly nuclear morphology and their positions within the osseus structures in regions that represent the ossification centers of their respective bones (Rugh 1968).

In addition to differentiating bone cells, several other cell types also stain for UNC119 and UNC119B at E14.5. Large cells within several follicles of the vibrissa (the

whisker primordia) display strong nuclear staining. Granular staining of non-hepatic cells is also observed in the developing liver. This staining is found only in certain large cells dispersed throughout the organ that have a nuclear morphology. Although the application of various discriminating stains to adjacent sections of embryonic tissue would be required to positively identify expressing cell types, they are suspected to be hematopoeitic in nature based on their morphology and pattern of dispersion (Zalik, personal communication). Furthermore, the liver is the principal hematopoeitic centre in mice during the E12.5-E16.5 gestational period (Rugh 1968), and UNC119 has recently been isolated from human eosinophils, which are also hematopoeitic cells (Cen *et al.* 2003).

The cells of the trigeminal ganglion (cranial nerve V) are also stained at E14.5, although in these cells of striated morphology, staining is much more diffuse. Unlike the cell bodies of the ganglion itself, the 3 nerves that emanate from it do not show any evidence of UNC119 expression.

Finally, strong staining of a few large cells dispersed throughout the intestinal mucosa (suspected to be endocrine cells) is also observed. However, such cells are often indiscriminantly stained by ISH because they non-specifically bind RNA (www.novocastra.co.uk/data/protocol/protocol.htm).

Although clear differences between the antisense and sense stained sections were apparent in this analysis, further experiments using additional controls have been performed to rule out the possibility of non-specific staining and to eliminate the possibility of cross-reaction between UNC119 and UNC119B. These experiments were conducted because all of the cell types stained by both probes were similar and are frequent sites of non-specific staining in non-radioactive ISH techniques (Fox, personal communication; Musat-Marcu and Nation, personal communication; www.novocastra.co.uk/data/protocol/protocol.htm). Furthermore, all staining was relatively weak, and at least some staining was observed in the same cells on the sense controls. Unfortunately, duplicate repetitions of the procedure using shorter, highly specific probes and endogenous peroxidase blockers in all wash steps have failed to yield any signal above background (data not shown), even under conditions of reduced

stringency. The significance of the initial results demonstrated in figures 18 and 19 are therefore highly questionable.

3.5 Production of α-UNC119 and α-UNC119B antibodies

Although the sites of gene expression may be a useful indication of cell types and tissues in which protein function is important, areas of protein localization and action may differ significantly due to cellular regulation of protein translation, post-translational modifications, or protein transport. Localization of UNC119 and UNC119B proteins has therefore been examined directly using Western analysis and immuno-chemical techniques.

Both techniques depend on the availability of antibodies specific to each protein. There is a previous report of polyclonal α-UNC119 antibodies (Higashide *et al.* 1998), however they are only available in extremely limited amounts that are insufficient for detailed analysis. α-UNC119 and α-UNC119B antibody generation and optimization was, therefore, a major focus of this study. As described below, this process involved production of large amounts of GST-tagged protein in a bacterial system, purification of the protein from the bacterial host, injection of purified protein into rabbits to induce a specific immune reaction, periodic titre analysis of the sera, and sera collection.

A) Expression of GST-fusion proteins in a bacterial system

Prior to this project, an expression construct (ufp-5) of the full-length human UNC119 coding region fused to glutathione S-transferase (GST) protein tag had been made (Maduro 1998). Production of this fusion protein in a bacterial system demonstrated that it migrated at ~60K and was approximately 50% soluble (Maguire, personal communication). Ufp-5 was transformed into a bacterial host, and specific induction of the fusion protein was confirmed prior to large scale protein purification. As shown in figure 20A, specific induction of a protein consistent with previous results was observed.

For UNC119B, an EST clone containing the entire known open reading frame of the predicted protein was obtained (gi:1165809), and its insert was cloned into an

appropriate pGEX expression vector. The cloning site was sequenced to confirm that the UNC119B and GST protein tag were fused in the same reading frame. As shown in figure 20B, strong specific induction in a bacterial host of a ~70K protein (expected to correspond to UNC119B-GST) was observed.

B) Purification of full-length GST-tagged UNC119 and UNC119B proteins

Several large scale protein inductions of bacterial strains containing each expression construct were performed, and large amounts of each fusion protein were purified. As shown in figure 21, the purified proteins have been examined for size and purity on a separating gel. Based on the assumption that the different proteins bind the same amount of dye per mass unit, their concentrations were estimated by comparison to known concentrations of bovine serum albumin protein (BSA). The UNC119-GST yield has been roughly quantitated to 0.2 μg/μl, while the UNC119B-GST yield is much higher (~34μg/μl). Qualitative examination of the protein gels demonstrates that the largest, brightest bands (shown in figure 21) migrate at the expected size for the full-length fusion proteins when compared to the induced bands demonstrated in figure 20. The smaller, fainter bands observed may result from partial protein degradation.

C) Antibody specificity against recombinant protein

Immune reactions were induced against each purified protein in 2 rabbits each, and boosts were continued until Western analysis of test bleeds indicated that titre saturation was reached. Complete sera collections following rabbit exsanguinations were preserved, and each antiserum recognizes the purified protein against which it was induced (figure 22).

D) Antibody specificity against retinal protein extracts

To determine whether the antisera are specific for the native forms of UNC119 and UNC119B, Western blotting against protein extracts from bovine retina was performed to test each antiserum. Retinal protein was selected because it is known to be a rich source of UNC119. It was not known, however, whether cows have an UNC119B gene, and if so, whether it is expressed in retina. Although the antisera were raised

against the human proteins, bovine retinas were available, and, at least for UNC119, previous studies indicated that antisera raised against it should cross-react in mammals (Higashide *et al.* 1998).

As demonstrated in figure 23A, the serum obtained from each α–UNC119 rabbit specifically recognizes a single ~33K product, the expected and published UNC119 protein size (Higashide *et al.* 1998). This detection is lost when exogenous UNC119-GST fusion protein is pre-absorbed to either antiserum, demonstrating that the protein recognized corresponds to native UNC119. Pre-absorption of the α–UNC119 antisera with similar concentrations of UNC119B-GST was also performed to confirm that the α–UNC119 antisera do not cross-react with this closely related protein. As shown in figure 23B, exogenous UNC119B-GST protein does not attenuate any signal detected by either antiserum. At least in retinal extracts, both α–UNC119 antisera therefore appear to be specific for UNC119.

Initial testing of the α–UNC119B 11-5 and 11-6 antisera against retinal protein resulted in detection of multiple bands and high levels of background. Each antiserum was therefore purified against membrane strips containing purified UNC119B-GST protein. As shown in figure 24A, the predominantly detected product using strip purified 11-5 antiserum migrates at ~55K and is specifically out-competed by UNC119B-GST. Similar to UNC119, UNC119B is therefore produced in retinal tissue. Other bands are also detected by purified 11-5 (figure 24B), however they are only visible after prolonged exposures, suggesting that the α–UNC119B antiserum faintly recognizes other proteins.

To further confirm that the most strongly detected protein corresponds to native UNC119B, retinal protein and UNC119B-GST recombinant protein were size fractionated in parallel. As expected, the detected band in the retinal extract migrates faster than the recombinant protein (figure 25A). However, the bands appear to have no more than a 5-10K difference in migration even though the GST moiety is 28.4 KDa and migrates at a rate consistent with its molecular weight. Both α–UNC119 and α–UNC119B antisera were also applied to parallel aliquots of retinal protein to confirm that the detected product of each antiserum is distinct (figure 25B), and no cross-reaction to the other protein by either antiserum was apparent even upon long exposures.

E) Antibody examination against brain extracts

Because UNC119 is reported to have at least some expression in brain at the transcript level (Swanson *et al.* 1998), protein was extracted from adult bovine brain and tested for the presence of UNC119.

Unlike the single band detected in retinal extracts, both α–UNC119 antisera faintly recognize several products in brain protein extracts. Each antiserum was pre-absorbed with high concentrations of exogenous protein to determine whether the larger products are out-competed by UNC119, and, as demonstrated in figure 26, none of these bands are clearly attenuated by pre-absorption with UNC119-GST.

However, multiple repetitions of this experiment have generated inconsistent results that are likely related to the high levels of background associated with the long exposures required to detect any signal. Furthermore, even though both OF3 and OF5 detect a product that migrates at a rate close to the one prominent in retinal extracts, the product detected by OF3 appears to be slightly larger than the one detected by OF5, suggesting that only one of the 2 antisera (if any) recognizes UNC119 in adult brain protein. Although no clear results were obtained from these experiments, further analysis was discontinued in favor of α–UNC119 characterization and optimization using tissues more specific to the objectives of this project.

F) Antibody examination against mouse extracts

The substrate used for the UNC119 immuno-localization portion of this project is tissue from mouse embryos, therefore each antiserum was Western blotted against embryonic mouse protein to test this extract for the presence of UNC119 and UNC119B. Similar pre-absorption experiments were also conducted to confirm the specificities of each antiserum in this tissue.

Similar to the results observed in brain extracts, both α–UNC119 antisera detect larger products than are expected for native UNC119 protein. As demonstrated in figure 27A, a ~70K band is detected by OF3 antiserum that appears to be specifically out-competed by UNC119-GST. A smaller product that migrates at a rate consistent with the known UNC119 size is also detected, however it does not appear to be specifically out-competed. OF5 antiserum faintly recognizes several larger products non-specifically. It

does not appear to detect the expected ~33K UNC119 band, even following saturating exposures (demonstrated in figure 27B).

Because several bands are detected by each α–UNC119 antiserum, the purification protocol previously adopted for α–UNC119B was performed against OF3 and OF5 to remove non-specific epitopes and increase their specificities in the tissue of interest. As expected, the purified antibodies specifically detect the protein in retinal extracts as had been observed prior to purification (data not shown). However, because purified OF5 detection of the retinal product was extremely faint, and multiple applications of this antibody against mouse protein failed to yield any significant signal detection (data not shown), only purified OF3 has been further characterized.

Purified OF3 strongly and specifically detects a single band of ~70K in E13.5-14.5 mouse protein, as demonstrated in figure 28A. Pre-absorption with exogenous UNC119-GST reduced the intensity of the ~70K band compared to untreated antiserum. To assess the degree of attenuation by UNC119 pre-absorption, scans of the Western blot were analyzed by densitometry using ImageJ software (Rasband 2002). As depicted in figure 28B, UNC119 pre-absorption of the purified antibody significantly attenuates the detected signal, while pre-absorption with neither UNC119B nor with an unrelated protein results in any attenuation. The relative differences in signal intensity between each treatment have been quantified by comparing the surface areas under each signal peak. As shown in Table 2, UNC119 pre-absorption results in a 72% reduction of signal intensity. Long exposures of the same membrane do reveal extremely faint detection of numerous products by the purified antiserum (data not shown), however the other products observed prior to purification (shown in figure 28A, lane 2) are no longer present.

Given the different migration rates of the products detected by purified OF3 in the various protein extracts examined, parallel immuno-precipitations using purified OF3 against several extracts were performed to further characterize and compare all products recognized by the antibody. As demonstrated in figure 29A, several proteins from each extract were eluted from α–UNC119 (purified from OF3) conjugated beads. Only the 3 strongest bands in each elution, when overlayed, are recognized by the purified antiserum following Western blotting (figure 29B). All products recognized by the antibody are

present in the eluted fractions from all extracts (brain, mouse, retina) examined. The ~55K product is expected to correspond to the immunoglobin (IgG) heavy chain of the antibody itself, and the larger and smaller ones migrate at rates consistent with the specifically out-competed bands in the mouse and retina extracts respectively. A larger scale pull-down and subsequent analysis of the 3 products by mass spectrometry could be performed to positively identify the products.

The presence of UNC119B protein in mouse embryo has also been tested using α–UNC119B purified from 11-5 antiserum. Although long exposures of Western blots are associated with high levels of background, they were necessary to detect UNC119B in mouse extract. Initial short exposures suggested that the un-purified serum detects a band that migrates at a different rate than the one recognized by the purified antibody (shown in figure 30A). Long exposures, however, demonstrate that α–UNC119B recognizes 5 different products that include those observed initially, although only 2 of the detected bands are specifically out-competed by pre-absorption with UNC119B-GST (shown in figure 30B). One of these out-competed products migrates at ~55K, a rate consistent with the out-competed product observed in retinal extracts. Purified α–UNC119B therefore appears to specifically recognize 2 products (and bind 3 others) in mouse embryo extracts.

Given the detection of multiple bands by the purified 11-5 antiserum, further characterization was performed (as described for α-UNC119) by immuno-precipitating proteins from various protein extracts using α–UNC119B. Although the concentration of eluted proteins was not very high, 2 different products were faintly visible in each extract (data not shown). As demonstrated in figure 31, both eluted products (~80K and ~55K) are present in each extract, and they are recognized by purified α–UNC119B. Mass spectrometry analysis (shown in Appendix 5) confirms that the smaller product is the antibody's heavy chain immunoglobin (IgG).

However, given that the predicted UNC119B protein migrates at approximately the same rate as the heavy chain, it is possibile that the ~55K band corresponds to both products. Unfortunately, sufficient yields of the larger product have not been obtainable for analysis.

3.6 UNC119 and UNC119B protein localization in mouse embryos

Although Western analysis has established the presence of UNC119 and UNC119B proteins in the mouse embryo, antibody staining of histological tissues was performed to characterize the spatial distribution of the proteins.

Initial trials of UNC119 immuno-histochemical localization were conducted using un-purified OF3 and OF5 antisera on frozen tissue sections of E11.5-14.5 mouse embryos. With the generous help of Dr. Putman, the optimal antigen unmasking reagent and sera titre for UNC119 staining of frozen mouse embryo tissue was determined. By comparing the staining patterns of nearly serial sections probed with pre-immune versus immune antiserum, some differential staining was observed. As demonstrated in figure 32, staining of rows of cells dorsal and anterior to the developing vertebrae (32B), and of small clusters of cells immediately dorsal of each spinal segment (32D) is predominant, with other signal seen elsewhere. The small clusters of cells may correspond to the nerves that emanate from each spinal segment (Bagnall, personal communication). However, even though these early results were promising, further analysis was performed using only purified OF3 antiserum because the un-purified antisera reacted with several bands against mouse protein on Western blots. Furthermore, good quality frozen tissue sections were never achieved.

Optimization of antigen unmasking conditions for sections of paraffin embedded specimens have since been determined, and α–UNC119 (purified from OF3) and α–UNC119B (purified from 11-5) were tested for specificity under different conditions to determine those that generate the strongest signals with the least non-specific background staining. Initial trials of UNC119 and UNC119B staining using roughly staged (E12.5-E14.5), poor quality tissue sections (data not shown) demonstrated that both proteins appear to be widely distributed at low levels throughout the embryo with some enrichment in the neural tube. However, at high magnification, the staining appeared faint and diffuse, a characteristic typical of non-specific antibody binding.

Analysis of commercial E8.5, E10.5 and E14.5 sections was therefore performed after antibody pre-absorption to differentiate antibody binding to UNC119 from binding to other products. As demonstrated in figure 33A and 33B, no detection of UNC119 is

apparent at E8.5. At E10.5 and E14.5 (figures 33C-F), however, low levels of signal are detected throughout the embryo compared to the pre-absorbed controls. However, close examination of stained sections indicates that staining is extremely diffuse, and even though the pre-absorbed control largely attenuates the observed signal, it is possible that the antibody also binds to other proteins. The pre-absorption could simply lower overall signal detection by reducing the concentration of antibody available for indiscriminant binding to any protein.

The specificity of the α–UNC119B antiserum for the UNC119B protein is even more questionable, because no significant signal attenuation is discernible between purified 11-5 treated sections and parallel pre-absorbed controls. Further controls have therefore been conducted to better characterize the purified antiserum, as demonstrated in figure 34C. Essentially no staining is observed on duplicate sections stained with high concentrations of pre-immune serum that was purified using the same protocol as the immune serum, and, as expected, there is also no visible signal on sections treated with non-specific rabbit immunoglobin (IgG) (data not shown). However, staining using un-purified pre-immune serum (figure 34B) yields a similar staining pattern to the one observed for purified 11-5 (figure 34A), although it is slightly fainter than the one observed for the purified antiserum at an equivalent titre. The detected signal observed following application of purified antiserum is present irrespective of whether or not it has been pre-absorbed with UNC119B-GST, indicating that the detection is not likely specific.

3.7 PC-12 cell culture preparation

Another aim of this project was to determine the sub-cellular localization of UNC119 and UNC119B to gain more insight into their biochemical functions. To begin to address this question, a cell culture was set up with the generous support of Dr. McDermid and other members of her lab.

The initial objective of the study was to test the rat pheochromocytoma PC-12 cell line for the presence of each protein in both un-differentiated and neuronally differentiated (NGF induced) cultures. If present in the cells before and after induction,

the sub-cellular localization of the proteins could then be monitored throughout growth and differentiation.

Despite multiple losses of culture due to difficulties involving bacterial contaminants, enough cells of a contaminated un-induced culture were harvested for protein extraction, and this extract was tested for the presence of UNC119 and UNC119B by Western analysis. As shown in figure 35, PC-12 cells produce significant levels of both UNC119 and UNC119B protein in the un-induced state. Upon short exposure, purified α–UNC119 and α–UNC119B each detect single bands (~70K and ~95K respectively). Faint levels of other proteins are observed, although they are only visible following saturating exposures, indicating that the purified antibodies should predominantly stain only the proteins against which they were raised if used for immuno-staining of plated cells. Although the tested extract does contain contaminants, it is extremely unlikely that the proteins detected were bacterially derived because no members of the UNC–119 protein family have ever been identified in prokaryotes.

Based on the presence of UNC119 and UNC119B and the specific recognition of each product in cultured PC-12 cells, future immuno-localization experiments are expected to be highly successful in determining the sub-cellular distribution of these proteins before, during, and following axonal outgrowth and differentiation.

3.8 Identification of UNC119 interacting proteins

The final objective of this thesis was to identify proteins that physically interact with UNC119, because this is critical to understanding the functional mechanisms of the protein's biochemical properties. Both genetic and biochemical tools are available and have been attempted to identify such interactions.

Prior to this M.Sc. thesis work, a genetic approach had been adopted to search for UNC119-interacting proteins using the yeast two-hybrid system. A human fetal brain cDNA library was screened to saturation, and although over 40 candidate interacting proteins had been identified, all were systematically ruled out as false positives. It is now known that the quality of the library screened was poor, therefore repetition of the screen using a different library could potentially identify authentic UNC119-interacting proteins.

A two-hybrid screen searching for UNC119B-interacting proteins has never been attempted and could also yield interesting results.

Given the difficulties and limitations encountered using the two-hybrid system, this study has used 2 biochemical approaches to search for interacting proteins. Preliminary GST pull-down and co-immuno-precipitation experiments for UNC119 were conducted once the required reagents were developed.

For the GST tag-based pull-down assay, UNC119-GST bound beads were incubated with non-denatured mouse protein lysate. A large number of proteins were eluted from UNC119-GST coated beads (shown in figure 36) that represent potential UNC119-interacting proteins. Given the large number of eluted products, however, experimental optimization using either salt or pH gradients of increasingly stringent washes would be required to identify specific UNC119 interacting proteins from the eluted fractions.

For the co-immuno-precipitation experiment, beads conjugated to un-purified α–UNC119 OF3 and OF5 antisera pulled down an identical pattern of proteins from mouse protein lysate. Because far fewer products were eluted from the antisera in this experiment compared to the GST tag-based assay, the co-precipitated proteins were resolved on a large gel, as demonstrated in figure 37. Several products were extracted from this gel and subsequently analysed by mass spectrometry. As expected, the strongest band corresponds to the immunoglobin (IgG) heavy chain of the antibody. Two other products were also analysed, and they correspond to mouse A-X actin (gi:309090) and to erythropoietin receptor protein (EpoR) (gi:193200). Protein summary reports for each analysis are depicted in Appendices 6, 7, and 8 respectively. As demonstrated in Appendix 8, the score for the match to EpoR is below the significance threshold assigned by the search parameters of the Mascot Search program (http://www.matrixscience.com/cgi/master_results.pl?file=../data/20020206/FamTSns.dat &REPTYPE=protein).

Discussion

The objective of the study presented was to determine the cellular function of UNC-119 proteins and to delinate the evolution of this highly conserved family. Specifically, this project has aimed to characterize the mammalian members of the UNC-119 family to resolve the dichotomy of initial studies pointing to extensive structural and functional conservation of UNC-119 across metazoans despite vast differences in their localization patterns.

4.1 UNC119B: A second mammalian UNC-119 homologue

A novel unc-119 homologue has been identified in mammals named 'unc119 (*C. elegans*) homolog B', and it is represented by the symbol UNC119B. This gene is structurally organized into 5 exons that span approximately 10 kilobases of human chromosome 12 at cytoband q24.31.

Although the human UNC119B transcript size is unknown because a full-length cDNA clone is not yet available, it is expected to be at least 3.6 Kb based on the full-length mouse transcript (gi:26349818) size of 3645 bases. The human and mouse UNC119B sequences are 86% identical at the nucleotide level over their coding regions. Although untranslated regions tend to be less conserved than coding regions, the genomic structural conservation between these orthologues (demonstrated in figures 12 and 14) suggests that the conservation extends over the entire transcripts. However, an UNC119B probe detects a 4.5 Kb product on Northern blots (figure 14), suggesting that an extension of the human UNC119B 3' untranslated region compared to the mouse orthologue is possible. Given that the most distal 1.3 Kb of the mouse UNC119B transcript is composed of sequence duplicated elsewhere in the genome, it is possible that the transcripts differ either in composition or length of this repeat sequence.

Another possibility that could account for the difference in transcript sizes is the existence of alternate transcripts of different lengths for mammalian UNC119B. Several partially overlapping incomplete clones for mouse have recently been identified in databases, possibly due to varying splice forms in this mammal. Furthermore, recent

alignments of rat UNC119B ESTs reveal 2 isoforms that differ only in their 3' untranslated regions. In contrast, only a single human UNC119B transcript is detected by Northern analysis, although transcripts of low abundance would not have been detected since the signal on the Northern blot was faint. Because there is a precedent for different levels of several transcripts for other family members such as UNC119, additional UNC119B transcripts, including one corresponding to the size of the full-length mouse clone, may exist.

Another feature of the predicted UNC119B transcript is that it is much larger than the UNC119 mRNA, which has a size of only ~1.5 Kb on Northern blots. This is unexpected given the conserved structural organization of unc-119 introns and exons across metazoans (Maduro *et al.* 2000). However, the coding region of UNC119B maps to the 5' portion of the transcript, therefore this difference in size may be accounted for by divergence in the 3'untranslated region, and this could, at least in part, be a consequence of a recent genomic insertion or modification of the repeated sequence identified at the 3' end of the mouse transcript.

The 5'untranslated region of human UNC119B also remains to be characterized, and its cloning will be required to unequivocally confirm the amino terminus of the protein. A predicted translation initiation site based on genomic sequence has been identified, although this "ATG" codon is not in an optimal context for initiation of translation in vertebrates (Kozak 1996). However, the perfect match of the 8 terminal amino acids predicted by translation from this site with the mouse UNC119B protein and with the predicted rat UNC119B protein suggests that it corresponds to the authentic human UNC119B 'ATG^{start}' site. Translation initiation from this site also predicts a size for UNC119B that is consistent with the sizes of the other UNC-119 family members, and the absence of any other in-frame upstream translation initiation sites further supports that initiation occurs at this site. It is unlikely that an additional 5' exon containing the 'ATG^{start}' site exists based on the conserved unc-119 5 exon structure. Furthermore, translation initiation from poor contexts has been observed for certain tightly regulated products that are encoded by stable transcripts. These include growth factors, transcription factors, or signal transduction components (Kozak 2000), and, as described

later, a role for UNC-119 in signal transduction has been proposed. It is therefore very likely that the 5' end of human UNC119B has been identified.

Domain analysis of the predicted UNC119B protein suggests that it is a non-globular protein with a cytoplasmic localization. Several structural features and post-translational modification sites have also been predicted, and their degrees of conservation across the UNC-119 family have been assessed. A possible reactive cysteine residue conserved in all members of the UNC-119 family may have suggested that a disulfide bridge is present, however such bonds are rarely observed in cytosolic proteins because of the high concentrations of reducing agents found in cell cytoplasms (Creighton 1988).

UNC119 and UNC119B also contain possible PEST motifs, which are hydrophilic sequences that target proteins for rapid destruction (Rogers *et al.* 1986). This could indicate tight regulation of both proteins *in vivo*, because targeted proteolysis is a cellular mechanism of control used to maintain low levels of these proteins in the cells (Rogers *et al.* 1986).

The least similar portions of UNC119 and UNC119B were examined to determine if they contain motifs that are specific to each protein, because differences in protein modifications could suggest that the homologues are differentially regulated and have non-overlapping roles. Several O-glycosylation sites are predicted for UNC119, while weak N-myristoylation signals are predicted in regions specific to UNC119B. However, because these predictions are relatively weak, and glycosylation occurs only on non-cytosolic proteins once they have entered the secretory pathway (Alberts *et al.* 1994), it is unlikely that the predicted motifs for either protein are sites of post-translational modifications.

Given that such predictions were not informative in predicting whether UNC119 and UNC119B have non-overlapping roles, functional conservation of the carboxyl domain of UNC119B should be confirmed. This could be assessed by determining if UNC119B is capable of rescuing *C. elegans unc-119* mutants, because transgenic expression of other family members (including UNC119) in *unc-119* worms is sufficient for full rescue of any obvious mutant phenotype.

Another reversible post-translational modification critical to the function of many proteins is phosphorylation. Several potential tyrosine phosphorylation sites have been identified that are conserved throughout several if not all members of the UNC-119 family. To verify the occurrence and locations of such modifications, radioactive phosphorylation incorporation assays and targeted mutagenesis could be performed to determine and compare the biochemical properties of UNC-119 proteins.

The predicted UNC119B protein sequence based on genomic analysis is not consistent with some experimental data. The UNC119B product detected in protein extracts migrates at a rate much slower on denaturating SDS-PAGE gels than is predicted based on its molecular weight from the primary protein sequence. Both UNC119-GST and UNC119B-GST fusion proteins migrate at fairly similar positions (demonstrated in figure 21), as would be expected for proteins closely related in size and composition. Their theoretical molecular weights calculated using their amino acid sequences are 55.3 KDa and 55.7 KDa respectively, yet both migrate at larger apparent mobilities (~60K and ~70K respectively) on denaturing gels. Because this has been observed for other members of the UNC-119 family (Maguire, personal communication; Manning, personal communication), and migration rates of proteins are significantly impacted by factors other than their amino acid lengths (Sambrook and Russell 2001), this is not entirely unexpected for UNC119B-GST.

Despite the fairly similar apparent molecular weights for the fusion proteins, the migration rates of the native forms of UNC119 and UNC119B in protein extracts contrast sharply. The UNC119B protein migrates at a molecular weight consistent with a protein nearly twice the size of UNC119, as shown in figure 25B. The apparent mobility of the pGEX-2T GST tag (Pharmacia Biotech) on denaturing gels corresponds to its theoretical molecular weight of 28.4 KDa. Based on the absence of this GST moiety, the observed ~33K migration of UNC119 is therefore consistent with the expected migration of the native protein (~32 KDa). For UNC119B, however, its native form would be expected to migrate at ~43K based on removal of the tag and addition of the 8 predicted 5' amino acids, yet its observed migration is much slower, at ~65K.

Properties affecting protein migration such as charge could explain the small difference in migration of UNC119B-GST compared to UNC119-GST, however this

does not explain the large difference in migration of native UNC119B compared to its expected size. In fact, the native protein in bovine retinal extracts migrates only slightly faster than the tagged fusion protein (~60K versus ~65K, shown in figure 25A) even though the fusion protein should be 218 amino acids longer. The average molecular weight of a single amino acid is 0.12 KDa, therefore the native form would have been expected to migrate ~26K faster. This discrepancy indicates that either the UNC119B protein is larger than predicted, that it has alternate isoforms, or that it undergoes post-translational modifications that do not occur in bacterial systems, the source of the fusion protein.

A potential modification that could explain this increase in molecular weight to almost twice the expected size is the formation of SDS insoluble dityrosine cross-links resulting in the formation of covalently bound protein oligomers. Such cross-links are induced by exposure to nitrating agents commonly found in the elevated oxidative environment of neuronal cells, and they have been observed for other proteins with similar characteristics to UNC119B, such as α-synuclein (Souza *et al.* 2000). Both are small proteins that are mainly cytoplasmic with α-helices in their secondary structures, and each contains several tyrosine residues. They also appear to associate, at least in some cases, with vesicles in the pre-synaptic terminal of neuronal cells (Davidson *et al.* 1998; Higashide *et al.* 1998).

Tyrosine nitration is a very selective process that is dependent on the local environment of the tyrosine residues (Ischiropoulos 1998), which could potentially explain why the highly related UNC119 protein migrates close to its expected size in retinal extracts while UNC119B does not. Cross-linking activity, while predominant for α-synuclein, is even absent in the highly related β-synuclein (Souza *et al.* 1999). There are many critical prerequisites for nitration, such as the absence of sterically hindering large amino acid side chains or reactive thiols and the presence of negatively charged residues in the vicinity of the tyrosine residues (Souza *et al.* 1999). Manual examination of UNC119B sequence indicates that tyrosine 142 is conserved in UNC119, is located near a very UNC119B-specific glutamic-acid rich charged domain, and it is not surrounded by hindering reactive cysteine residues (figure 38). Whether this modification occurs at tyrosine 142 of UNC119B remains to be determined.

Searches for motif similarities of UNC119B to other protein families have also been performed. Neither MaxHom nor ProDom, multiple sequence alignment programs based on remote similarities of protein domains, reveal any significant matches of UNC119B domains to proteins other than UNC-119 or PDEδ family members (Sander and Schneider 1991; Corpet and Kahn 1998). This is consistent with previous observations for UNC119.

However, a recent publication reveals that UNC119 contains Src homology (SH2 and SH3) protein binding motifs (Cen *et al.* 2003). This group showed expression of UNC119 in eosinophils, and they have demonstrated that it is involved in signal transduction via protein-protein interactions. The interactions between UNC119 and other proteins are mediated through a highly conserved carboxyl SH2 domain spanning residues 191-198 (corresponding to residues 202-209 of UNC119B) and through an amino terminal SH3 domain spanning residues 56-62. As shown in figure 38, the SH3 domain is not conserved; it is specific to mammalian UNC119. A second incomplete UNC119-specific SH3 motif that spans residues 26-32 has also been identified. Although $SH3^{26\text{-}32}$ contains the canonical "PXXP" SH3 binding motif (2 proline residues separated by any 2 amino acids), it is not optimal because it lacks the flanking arginine (R) residue of the class I type SH3 "RXXPXXP" motif that is critical for specificity of binding to and activation of the Src class of tyrosine kinases (SrcTKs) (Alexandropoulos *et al.* 1995). Tyrosine phosphorylation sites for SrcTKs are also present in UNC119 at residues 166 and 194 (corresponding to tyrosines 177 and 205 for UNC119B), both of which are highly conserved. Tyr^{166} is conserved in all known vertebrate UNC-119 proteins, while Tyr^{194}, the SH2 domain consensus tyrosine phosphorylation site for SrcTKs (Songyang *et al.* 1993), is conserved in all UNC-119 proteins. Both of these potential UNC-119 phosphorylation sites were also predicted as probable modification sites in the UNC119B domain analysis.

A highly conserved SH2 Src kinase protein binding domain with its associated tyrosine phosphorylation site suggests that this represents a functional domain of UNC-119 and provides strong evidence for it having a role in regulation of Src type protein kinases. Src kinases are critical components of multiple signal transduction cascades. The exclusiveness of the SH3 domain for UNC119 indicates that it may be important for

specificity of regulation to a particular signal, and it could explain how different UNC-119 proteins could have non-redundant biological roles despite a conserved biochemical function.

4.2 Vertebrate UNC-119 expression patterns

One of the major objectives of this study was to characterize the temporal and spatial patterns of mammalian unc-119 expression. Confirming the recently published non-retinal expression of UNC119 (in myeloid and lymphoid cells) (Cen *et al.* 2003), this study provides evidence for widespread non-retinal expression of both UNC119 and UNC119B in mammals. This study is also unique in that it is the first to examine expression of UNC119B, and of UNC119 during development.

Widespread expression of UNC119B was immediately apparent based on the initial identification of several corresponding ESTs extracted from multiple types of tissues during various stages of the life cycle. Northern analyses of human fetal tissues for both UNC119 and UNC119B demonstrate that both genes are expressed long before the differentiation of retinal tissues in mammals takes place. Although tissue extractions for Northern blots are crude and each organ extract contains many cell types, it is apparent that both genes are expressed in multiple locations in the fetus. The adult Northern analysis of UNC119 performed during this study is consistent with previous observations, because no expression has been detected in any adult non-retinal tissues on Northern blots examined by myself or by others (Higashide *et al.* 1998). Similar to UNC119, UNC119B expression during adult stages was not detected using this technique.

Characterization of pre-natal UNC119 and UNC119B expression via *in situ* hybridization of whole mouse embryos reveals that both genes are widely expressed in a non-uniform pattern. Although specific regions of probe concentration and tissue morphologies are not entirely consistent between trials for older embryos, in every situation the UNC119 and UNC119B antisense probes stain the entire embryos much more strongly than do unrelated probes or sense controls. This is clear even when controls are applied at high concentrations and developed for long periods of time. The

staining patterns observed for UNC119 and UNC119B can therefore be attributed to mRNA expression rather than to non-specific staining.

Based on this initial examination, it is impossible to determine whether UNC119 and UNC119B display similar or distinct expression patterns during development. The possibility of cross-reaction between UNC119 and UNC119B by each probe has not been ruled out, because the probes used for each homologue contain a region spanning approximately 150 nucleotides that is 75% identical. However, cross-reaction could be tested directly by probing of dot blots containing small amounts of un-labeled UNC119 and UNC119B transcripts. Another property of the probes used is that they were long (1.3 Kb and 1 Kb respectively), and this may account for some of the non-uniform staining observed in body cavities such as the otic canal or the fragile frontal lobes of the head. These structures are easily ruptured and trap fluid.

Higher resolution analysis of expression is also required to identify the specific cell types expressing each gene. Non-radioactive *in situ* hybridization using mouse thin tissue sections representative of various developmental stages was performed in an attempt to obtain such data, however no conclusive results were obtained.

An alternative approach to characterize the expression patterns of UNC119 and UNC119B would be the use of a radioactive *in situ* hybridization technique on tissue sections. Although this approach is more dangerous and best avoided when other alternatives are available, the use of radiolabeled riboprobes is more sensitive and is associated with reduced levels of non-specific staining. The use of very sensitive high energy isotopes allows for detection of faint levels of hybridization and for signal quantitation by phosphor imaging (http://www.molecularhistology.com/methods.htm#insitu). In contrast, less sensitive non-radioactive techniques are very useful for detection of abundant transcripts, however the faint widespread expression patterns typically observed for ubiquitously expressed genes are often too low to be discriminated from background staining with this method.

Given that widespread low level expression is observed in the UNC119 and UNC119B stained whole embryos, the radioactive approach could be a viable alternative to the method used for this project. There is a precedent for the detection of transcripts in whole embryos using less sensitive techniques than are required for detection in

histological sections, because the overlay of many cell layers in intact tissues make fainter levels of expression more visible (Fox, personal communication). Furthermore, this has been observed for *Pax1*, the antisense transcript used in parallel to UNC119 and UNC119B for this study. *Pax1* signal was detectable using the non-radioactive approach only in whole embryos, even though successful *Pax1 in situ* hybridizations of sections have been previously reported using a radioactive approach on mouse sections of various stages (Deutsch *et al.* 1988). High energy ^{35}S-labeled riboprobes had been used, and prolonged exposures of a full week had been necessary to detect *Pax1* transcript accumulation (Deutsch *et al.* 1988). This level of expression is unlikely to be discernible using non-radioactive probes.

Another advantage of a radioactive approach is a reduction in non-specific staining of many cell types. Cells that contain high endogenous levels of the same enzymes that are conjugated to the antibodies used in non-radioactive techniques (alkaline phosphatases) tend to react with the developing reagents used in the protocol, producing a non-specific positive signal. This is frequently observed in the endocrine cells of the intestinal mucosa and in the ossifying bone cells, 2 of the most prominently stained cell types observed during this study. No cell types are intrinsically radioactive, therefore signal development by autoradiography is not expected to be biased towards false detection of certain cell types.

RT-PCR is a completely different approach that could also be used to further characterize UNC119 and UNC119B expression and to resolve whether or not their expression patterns are identical. Although this would require collection of additional embryos during different stages of development, this sensitive assay is a conclusive and quantitative method to discern the levels of expression at different stages, and it has been used successfully to demonstrate adult expression of UNC119 in eosinophils (Cen *et al.* 2003). Individual cell types cannot be distinguished using this method, although dissection of tissues from different organs of older staged embryos provides a more detailed profile of expression sites than is possible by Northern analysis or whole mount *in situ* hybridization.

4.3 Distribution of mammalian UNC-119 proteins

A preliminary temporal and spatial characterization of the UNC119 protein distribution pattern clearly demonstrates that it is present during several stages in non-retinal tissue. Using polyclonal antibodies raised against UNC119, the presence of this protein is demonstrated in mouse embryo, adult bovine brain, and un-differentiated rat pheochromocytoma PC-12 cells by Western analysis. These results are consistent with recent evidence of protein localization of UNC119 in several (non-retinal) adult mouse tissues including heart, liver, lung, and spleen (Cen *et al.* 2003). Specificity of the antiserum used in the analysis against UNC119 was confirmed (figure 23) by repetition of previous reports of a single ~33K detected product in retinal protein extracts (Higashide *et al.* 1998).

A similar examination of UNC119B distribution provides the first conclusive evidence of translation of a second mammalian UNC-119 homologue. Confirmation of specificity of polyclonal antiserum against this protein was demonstrated by detection of a single product in retinal and PC-12 protein extracts. The absence of cross-reaction between both mammalian UNC-119 proteins for each antiserum was also demonstrated (figure 25B), indicating that the product detected by α-UNC119B does not correspond to UNC119. Western analyses against several protein extracts demonstrate both retinal and non-retinal localization of this protein. Specifically, significant levels of UNC119B are observed in bovine retina, mouse embryo, and rat pheochromocytoma PC-12 cells. Its presence has not been tested against adult bovine brain.

An unexpected characteristic that was observed for both UNC119 and UNC119B is the large difference in size between the native products detected in the various extracts examined. Although the apparent molecular weights of the proteins could vary in the different extracts because they were extracted from different organisms, the nearly identical (>90%) conservation of the proteins in length and in amino acid composition indicates that species-specific variations do not account for the huge discrepancies observed. In fact, as summarized in table 3, the observed migrations for both proteins appear to be nearly twice as large in non-retinal extracts (~70K and ~80K for UNC119 and UNC119B respectively) compared to the retinal ones (~33K and ~55K). Immuno-

precipitations using each antiserum against the various extracts indicate that both products are present in each extract, confirming that the observed size differences are not a consequence of species-specific variations. The extract-dependent preferential detection of each product therefore indicates that the temporal and spatial distributions of the smaller and larger species are different and specific.

Although such observations were initially unexpected, there is precedent for the existence of 2 different isoforms of UNC-119 proteins. Temporal Western analyses of Unc119 and Unc119b in zebrafish reveal the presence of a second higher molecular weight product in whole protein extracts that were extracted during very early embryonic stages. A smaller product consistent with the expected size of each product is also detected during these stages, although the smaller products continue to be detected throughout development while the larger ones are no longer detectable after approximately 22 hours post fertilization (Manning, personal communication).

Because an attempt to confirm the identities of the smaller and larger mammalian products by mass spectrometry was not successful, it is impossible to unequivocally determine the basis of their differences at this time. It is possible that either alternate isoforms for each protein exist, or that varying proportions of both proteins undergo post-translational modifications that are resistant to denaturing agents. Although only a single partial clone has been identified to date, the presence of a mouse embryo transcript (gi:26337100) with an extended second exon suggests that alternately spliced transcripts with a different open reading frame are produced at least in some tissues. The translation of this putative second open reading frame remains to be determined.

It is also possible that the larger products represent modified versions or complexed forms of the smaller ones. As described previously, domain analyses of both proteins reveal several potential modification sites for UNC119 and UNC119B. Given that larger derivatives of both proteins are detectable in similar patterns, it is likely that any responsible modification for the increased apparent molecular weights is one that is conserved between the proteins. Because the observed migration rates of the larger derivatives of both proteins are nearly twice the size of the smaller ones, protein dimerization of a certain cellular fraction of each protein is likely. As described previously, SDS-PAGE resistant covalent dimerization through the formation of

dityrosine crosslinks is possible. Although the potential modification site appears to be more optimal for UNC119B than for UNC119, the possibility of such a modification for UNC119 has not been ruled out. The strong dependence of this type of modification on the specific conditions of the cellular environment surrounding the protein could also potentially explain the consistent differences in the distributions of the larger and smaller derivatives in the various protein extracts examined.

Similar to the expression studies, it is not possible based on this study to determine whether UNC119 and UNC119B have identical, partially overlapping, or unique localization patterns. Although the immuno-histochemical staining experiments performed have yet to generate conclusive results, α-UNC119 and α-UNC119B staining of mouse sections indicates that a ubiquitous distribution of both proteins appears probable.

Unfortunately, the significance of the staining observed is unclear, because the patterns for both proteins are very diffuse. Although this might be expected for ubiquitously expressed proteins, general diffuse staining patterns are also typical characteristics of non-specifically stained tissue. For certain antibodies, their specific recognition of the antigen against which they were raised is completely dependent on the chemical used to fix the tissues on which they are applied, regardless of any antigen unmasking reagents used to treat the tissue prior to immuno-staining. There is no universal fixative that is ideal for the demonstration of all antigens, and some antigens will not survive even moderate amounts of aldehyde fixation (http://www.ihcworld.com/introduction.htm). Furthermore, general non-specific staining of formaldehyde and paraformaldehyde tissues has been observed during immuno-histochemical staining trials using antisera raised against other UNC-119 proteins from both invertebrate and vertebrate members of this protein family (Hawley 2002; Manning, personal communication). The *Danio rerio* α-Unc119 antiserum has recently been demonstrated to show clear specific detection of Unc119 in trichloroacetic acid (TCA) fixed embryos despite indiscriminant staining on paraformaldehyde-fixed tissues (the fixative used for all mammalian sections examined) (Manning, personal communication). It is therefore likely that future analysis using alternatively fixed tissues will yield significant staining of UNC119 and UNC119B on mouse tissues.

4.4 UNC119 interacting proteins

Preliminary antibody pull-down experiments reveal that erythropoietin receptor protein (EpoR) and A-X actin co-precipitate with UNC119 from mouse embryo protein extracts. Although both proteins could potentially correspond to UNC119 interacting proteins, the experiment was performed using un-purified antisera, therefore its repetition using α-UNC119 purified from OF3 is necessary before any conclusions are drawn. Furthermore, the presence of UNC119 in the immuno-precipitate was not demonstrated directly by mass spectrometry, therefore the products identified in the elution cannot definitely be attributed to an association with UNC119.

However, even though the experiment performed was crude and intended to be a preliminary trial, the results obtained may have some significance. Size fractionation of eluates obtained from pull-down experiments using antisera obtained from both rabbits immunized against UNC119 display parallel patterns of co-precipitating proteins. Furthermore, interactions between UNC119 with both A-X actin and EpoR are plausible based on our current knowledge of UNC-119 proteins.

Although actin is a large protein with multiple binding partners (Ayscough 1998) and could be expected to bind bead matrix or precipitating proteins non-specifically, there is a precedent for a functional relationship between UNC119 and actin. Actins are highly conserved, ubiquitously expressed proteins that polymerize into structural filaments and are involved in various types of cell motility and contraction (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=Protein&list_uids=01351867&dopt=GenPept). Although there are multiple types of actin proteins in vertebrates that have varying functions (Hill and Gunning 1993), and no characterization of A-X actin has been published to date, sequence comparisons of actin transcripts suggest that A-X actin is either a beta or a gamma actin isoform. Of the 6 types of mammalian actins, beta and gamma are the cytoplasmic isoforms that are found in all non-muscle cells (Hill and Gunning 1993) and are critical components of the cytosketeton.

In addition, complementary genetic analysis in *Drosophila melanogaster* reveals a synergistic phenotype when a DmUNC-119 deficiency is crossed into a Profilin-deficient

chickadee mutant background, indicating that the proteins may be involved in a common functional pathway (Hawley 2002). Profilin is a cytoskeletal actin-binding protein (Carlsson *et al.* 1977) that promotes actin filament polymerization in motile cells in response to binding of tyrosine kinase signaling molecules (reviewed in Wills *et al.* 1999; Suter and Forscher 2000). This suggests that UNC119 may function in a complex with Profilin and actin in a pathway of actin cytoskeletal modeling.

An interaction between UNC119 and EpoR is also plausible even though the mass spectrometry analysis match to this protein is slightly below the assigned significance threshold. Recent reports clearly demonstrate an interaction between UNC119 and the interlukin receptor (IL-5) in eosinophils, and IL-5 and EpoR are closely related cell-surface receptor proteins that are members of the cytokine receptor superfamily (D'Andrea and Gonda 2000).

These receptors mediate proliferative, activation, and differentiation signals governing hematopoeisis (the production of myeloid and lymphoid cells) in response to cytokine growth factors (Geijsen *et al.* 2001). Upon cytokine ligand binding, the receptors elicit their diverse biological responses via activation of a number of intricately regulated signal transduction cascades (Rare and Reddy 2002). These cascades are mediated through the triggering of kinases associated with the intra-cellular domains of the receptors, and they result in a series of activating phosphorylation events (D'Andrea and Gonda 2000). Triggering the associated kinases (members of the Janus protein kinases) is the critical first step in cytokine receptor signaling (D'Andrea and Gonda 2000), and their activation results in several tyrosine phosphorylations on the receptor. The phosphorylation sites on the receptor proteins then become docking sites for Src tyrosine kinases and other adapter proteins via their SH2 domains (Geijsen *et al.* 2001; Fisher 2003), allowing these signaling intermediates to mediate the cytokine response by activating several cytosolic molecules (Cen *et al.* 2003). Because the UNC-119 proteins contain a conserved SH2 domain that includes its tyrosine phosphorylation site, their role as a receptor-associated protein is plausible.

If confirmed, interactions between UNC119 with A-X actin and EpoR would suggest that UNC119 may be an adapter molecule that functions in a complex linking Src tyrosine kinase signaling to actin cytoskeletal regulation. In addition to direct actin

binding, Profilin is known bind to members of the Enabled (Ena) protein family, which in turn bind the SH3 domains of Src tyrosine kinases (Gertler *et al.* 1995; Ahern-Djamali *et al.* 1999), providing another link between Src tyrosine kinase signaling and cytoskeletal modulation.

However, confirmation of the A-X actin and EpoR interactions will be critical before definite conclusions may be drawn. This is particularly true for EpoR, because although it is expressed as a protein that ranges from 66-78 KDa (Fisher 2003), the molecular weight of the product that co-precipitated with UNC119 was much smaller (only ~35K). The co-precipitated product could potentially represent only a fragment of EpoR. The peptide sequence obtained from mass spectrometry of this co-precipitated product is located in the N-terminal of EpoR (residues 6-25), which is part of its extra-cellular domain (Ketteler *et al.* 2002), very distant from the expected intra-cellular site of the UNC119-EpoR interaction. For the suspected SH2 domain interaction to have taken place, damage or dissociation of the EpoR protein following the co-precipitation experiment would have had to occur. Dissociated EpoR fragments would have been separated during the denaturating size fractionation on the protein gel, therefore other products in the eluate that remain to be analysed could correspond to the intra-cellular domain of EpoR. Western analysis of the eluted fraction using α-EpoR and α-actin antibodies would confirm the co-precipitation of these proteins with UNC119. Alternatively, the UNC119 protein-protein interactions could be confirmed using other techniques, such as directed co-immuno-precipitations or two-hybrid analyses.

4.5 Vertebrate UNC-119 family: General conclusions and speculations

The objective of this M.Sc. thesis project was to determine the cellular biological function of UNC-119 and to delineate the evolution of its protein family. Although this primary objective remains to be addressed, several new lines of evidence have been presented that suggest an important role for UNC-119 proteins in intra-cellular signal transduction.

A second highly related mammalian UNC-119 homologue, UNC119B, has been identified, and widespread non-retinal expression of UNC119 and UNC119B has been

clearly demonstrated during pre-natal development. The observed expression also transcends the nervous system. Both proteins are translated in several cell types during different stages of development, and analysis of their structural properties indicates that they contain a highly conserved Src tyrosine kinase homology 2 (SH2) domain and likely undergo tyrosine phosphorylation at this site. They may also be subjected to other modifications important for maintenance of their structural conformation, and to tight regulation of their levels via signals that target them to proteolytic degradation machinery. All of these features point to a role for the proteins as important adapter proteins in Src tyrosine kinase receptor-induced signaling cascades.

A proposal by others that UNC119 plays an important role in the retina (Kobayashi *et al.* 2003) is plausible. Based on comparisons of the functions of the highly related ARL proteins, an association between UNC119 with ARL1, ARL2, and ARL3 does postulate a role for UNC119 in membrane protein trafficking in synaptic vesicles and microtubule formation in photoreceptor axons. ARL proteins are associated with other proteins that are critical to a functional retina. For example, an interaction between ARL3 and the retinal disease gene retinitis pigmentosa 2 (RP2) has been documented, and the proteins have been ascribed a role in retinal microtubule regulation (Bartolini *et al.* 2002).

However, based on the evidence provided in this study and by others (Hawley 2002; Cen *et al.* 2003), cytoskeletal regulation in the retina likely represents only a subset of the effects of a more generally functioning UNC119 protein. Evidence from other ARL2 orthologues that have universal roles in microtubule dynamics (Kobayashi *et al.* 2003), and from the distant relative of UNC119, PDEδ, that has a universal signaling role (despite its enrichment in photoreceptor cells (Li *et al.* 1998)) demonstrates that UNC119 has a general role in cytoskeletal maintenance as opposed to a photoreceptor specific function.

The recent discovery that UNC119 binds to the cytokine cell surface receptor IL-5 (Cen *et al.* 2003) and possibly to the highly related EpoR receptor (figure 37) supports a model for UNC-119 proteins as adapters in Src tyrosine kinase (SrcTK) signaling cascades. In addition to the demonstration that UNC119 interacts with the α subunit of the IL-5 receptor in eosinophils, Cen *et al.* have revealed additional insights into the roles

of UNC-119 proteins by determining the biological significance of this interaction. They have shown that the protein and receptor associate independently of the activation state of the receptor, and that the interacting fraction of the proteins increases upon stimulation by cytokines. That the physical association of UNC119 with IL-5 is limited to its α subunit is also significant, because this subunit is important for the specificity of the biological response of the particular receptor upon stimulation (Evans *et al.* 1999).

Receptor stimulation in eosinophils triggers a series of Src tyrosine kinase phosphorylation events, therefore direct binding of UNC119 SrcTKs has also been tested and demonstrated (Cen *et al.* 2003). UNC119 specifically binds to SH2 and SH3 domains of SrcTks (in a non-phosphotyrosine-dependent manner) in the myeloid and lymphoid cells studied. Other non-Src type kinases containing the same SH motifs (Itk, Grb2, Shc) demonstrate no affinity for UNC119, confirming its specificity to SrcTKs.

Furthermore, UNC119 binding to the SH domains of Src kinases activates them and increases their catalytic activities in a dose-dependent manner. Eosinophil stimulation of IL-5 increases the fraction of UNC119 that is associated to the Lyn SrcTK and also promotes Lyn-mediated eosinophil survival and differentiation. Attenuation of UNC119 activity by addition of α-UNC119 antibody partially inhibits IL-5 eosinophil stimulation, confirming UNC119 as the first identified receptor-associated activator of SrcTKs (Cen *et al.* 2003).

Although eosinophils are unique types of terminally differentiated non-proliferating cells, it is likely that this function for UNC119 has broad biological significance. UNC119 exerts similar effects on SrcTKs in T cells (Gorska *et al.*, cited in Cen *et al.* 2003), and Src tyrosine signaling cascades are also important in non-hematopoeitic cells. Consistent with a role for UNC-119 during axonal outgrowth in invertebrates, increasing lines of evidence demonstrate that growth cone motility and guidance depends on dynamic regulation of the actin cytoskeleton and maintenance of membrane trafficking at their leading edge in response to extra-cellular guidance cues (Wills *et al.* 1999; Suter and Forscher 2000). Several studies point to the interpretation of these extra-celluar cues via complex signaling cascades involving Src tyrosine kinases. In particular, the requirement of the *Drosophila* Abl tyrosine signal transduction pathway for motor axon outgrowth has been well documented (reviewed in Wills *et al.* 1999), and

it elicits its cellular responses via SH domain interactions with the Enabled (Ena) protein (Gertler *et al.* 1995). Because Ena mediates its cytoskeletal effects through the actin-binding protein Profilin (Ahern-Djamali *et al.* 1999), and Profilin and UNC-119 deficiencies are synergistic (Hawley 2002), it is likely that UNC-119 is also an important adapter molecule in this Src tyrosine kinase mediated signal.

Tyrosine kinases play a role in communication between cells in multi-cellular organisms. These molecules are activated by a number of cell-surface receptors when they are bound by extra-cellular signaling ligands, and they mediate diverse cellular changes in response to ligand binding by activating multiple types of tyrosine phosphorylation signal cascades.

As shown in figure 39, UNC-119 appears to represent a novel class of signal transduction proteins whose role in different cell types or during different developmental stages depends on the downstream effectors that they activate and the specific responses that the effectors induce. The bipartite structure of UNC-119 into a highly conserved SH-containing domain and a poorly conserved N-terminal protein binding domain makes it an ideal candidate for an adapter molecule responsible for the specificity and transmission of SrcTK signaling. A conserved function for UNC-119 in activation of SrcTKs in response to receptor stimulation could convey the multiple types of tyrosine kinase mediated responses observed for the different UNC-119 proteins based on multiple lines of evidence. Downstream effectors that bind SH2 domains are differentially expressed in different cell types, and each could potentiate the unique cell and organism-specific effects observed based on which specific effector is expressed in a given UNC-119 containing cell. Furthermore, the different N-terminal binding capacities of each UNC-119 protein could further increase the specificity of cellular responses to signals by altering the overall composition of the SrcTK-activated protein complexes, suggesting a high likelihood for overlapping, non-redundant roles of the mammalian homologues, UNC119 and UNC119B. The UNC-119 proteins therefore appear to have evolved within an ancient highly conserved signal transduction pathway that has acquired several roles throughout metazoan evolution.

At this stage of analysis, it is not possible to resolve when the duplication event(s) of the unc-119 genes took place during the adaptive radiation of vertebrates. The total

numbers and timings of vertebrate duplication events remain to be determined, as is the level of temporal and functional overlap between the 2 highly related mammalian homologues, UNC119 and UNC119B. Determining their evolutionary relationships will give further insights into UNC-119 functions, because the probabilities of redundancies between homologous proteins vary greatly depending on their biological roles. Redundancy is widespread among complex metazoan developmental genes, although it tends to be rare among genes involved in general cellular maintenance (Cooke *et al.* 1997). Furthermore, apparently redundant genes often have combined emergent functions, and this is especially likely for underlying developmental genes that are expressed in a spatially restricted pattern (Cooke *et al.* 1997). Delineation of the precise relationships and functional overlaps between the UNC-119 proteins will therefore be critical to understand the full biological significance of the UNC-119 family.

Concluding Remarks

The dichotomy of initial studies pointing to extensive structural and functional conservation of the UNC-119 protein family across metazoans despite drastic differences in their localizations has been resolved. Two key observations resulting from this project explain how reports of a restricted localization of mammalian UNC119 in the adult retina could occur despite the extensive distribution of other UNC-119 proteins. Additional (non-retinal) sites of localization and function have been demonstrated for UNC119, and a second highly related mammalian homologue, UNC119B, has been identified. It will be interesting to conduct future studies to determine the precise relationship between UNC119 and UNC119B in both a functional and evolutionary context. Further characterization of these proteins is also likely to confirm their biochemical activities and allow the delineation of their precise cellular functions.

Table 1. Potential UNC-119 homologues in Fugu.

Fugu region	% matches to a *H. sapiens* homologue (out of 107 non-conserved a.a.)	Percentage of matches to	
		UNC119	UNC119B
1	48	10	**38**
2	52	8.5	**44**
3	37	**23**	15

a.a.= amino acid

Table 2. Densitometry analysis of α-UNC119 signal intensity on Western blots.

Lane	Area under curve mass (g)	Proportion of signal detected after pre-absorption (%)	% of signal reduceable by pre-absorption with UNC119-GST
PI	.0046	n/a	n/a
IS	.1384	n/a	89%
PS	.0526	100%	72%
UNC119	.0149	29%	n/a
UNC119B	.0903	170%	84%
BSA	.0486	92%	69%

Sample calculations:

Proportion of signal detected after pre-absorption (%):

- Purified OF3 antiserum was arbitrarily defined as 100% signal detected.
- The proportion of signal detected by purified OF3 following pre-absorption compared to non pre-absorbed purified OF3 was calculated as follows:

$$\frac{\text{area under pre-absorbed antiserum curve (g)}}{\text{area under non pre-absorbed purified OF3 curve (g)}} \times 100\%$$

% of signal reduceable by pre-absorotion with 0.11μg UNC119-GST per μl antiserum:

- The relative difference in signal intensity of UNC119 pre-absorbed antiserum compared to each other treatment was calculated as follows:

$$\left[1 - \frac{\text{area under UNC119 pre-absorbed antiserum curve (g)}}{\text{area under other curve (g)}}\right] \times 100\%$$

Table 3. Summary of the migration rates of the products detected by α-UNC119 and α-UNC119B antisera.

	Molecular Weight[t]	Cow Retina*	Rat Retina	Rat PC-12 cells	Mouse Embryo	Human Blood[◊]
UNC119	27.0	~33	33	~70	~70	37
UNC119B	28.1	~55	--	~95	~80	--
				?	~55	

t = Values displayed are in kilodaltons (KDa), calculated from cDNA sequences.

References

* = Inana *et al.* 1998.

◊ = Cen *et al.* 2003.

Figure 1. *C. elegans unc-119* mutants have a disorganized nervous system. A) The main components of the nematode nervous system: a large ring of bundled neurons near the head, dorsal and ventral nerve cords bundling neurons along the body length, and regular intervals of commissural neurons that extend along the circumference of the worm. B) Green fluorescent protein (GFP) expression in the ventral nerve cord of a wild-type worm shows regularly spaced commissural neurons and other longitudinal neurons. C) GFP expression in *unc-119* mutants (*unc119(ed3)*) demonstrates incorrect spacing of commissural and longitudinal neurons and extensively branched neurons. D) An α-GABA stained *unc-119* ventral nerve cord displays loose bundling of neurons and crossed or looped axons. Neuronal GFP expression was detected using the pF25B3.3:GFP reporter. Courtesy of W. Materi.

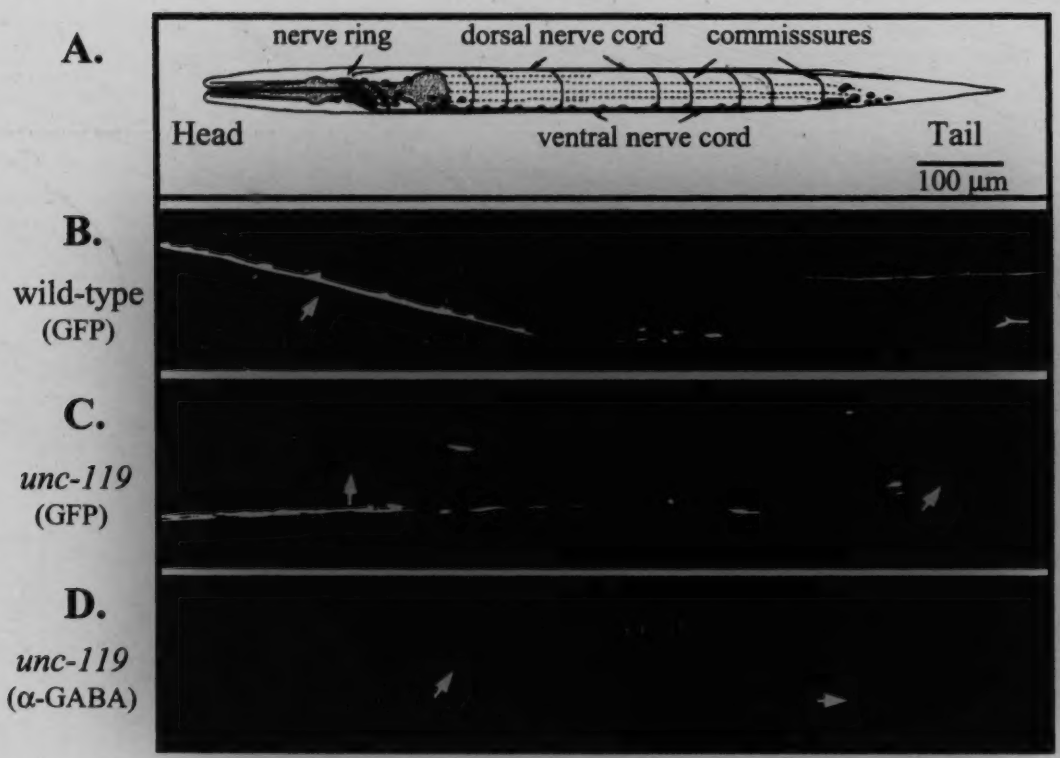

A.
nerve ring
dorsal nerve cord
commisssures
Head
ventral nerve cord
Tail
100 µm
B.
wild-type
(GFP)
C.
unc-119
(GFP)
D.
unc-119
(α-GABA)

Figure 2. UNC-119 is present in axons and cell bodies of *C. elegans* neurons. A) GFP expression of a P*unc-47*::GFP construct in GABA motor neuron cell bodies (arrows) and axons (**) of a wild-type worm. B) In the same wild-type individual, α-UNC-119 antibodies label multiple cell bodies (arrowhead and arrows) and axons in the left (*) and right (**) ventral nerve cords. C) GFP expression in a GABA motor neuron commissure (arrow) near the right ventral nerve cord (**) in a wild-type worm. D) α-UNC-119 antibodies label the axons in the right ventral nerve cord (**), the commissures seen in C (arrow), and the lateral nerve cord (arrowhead). From Knobel *et al.* 2001, figure 2.

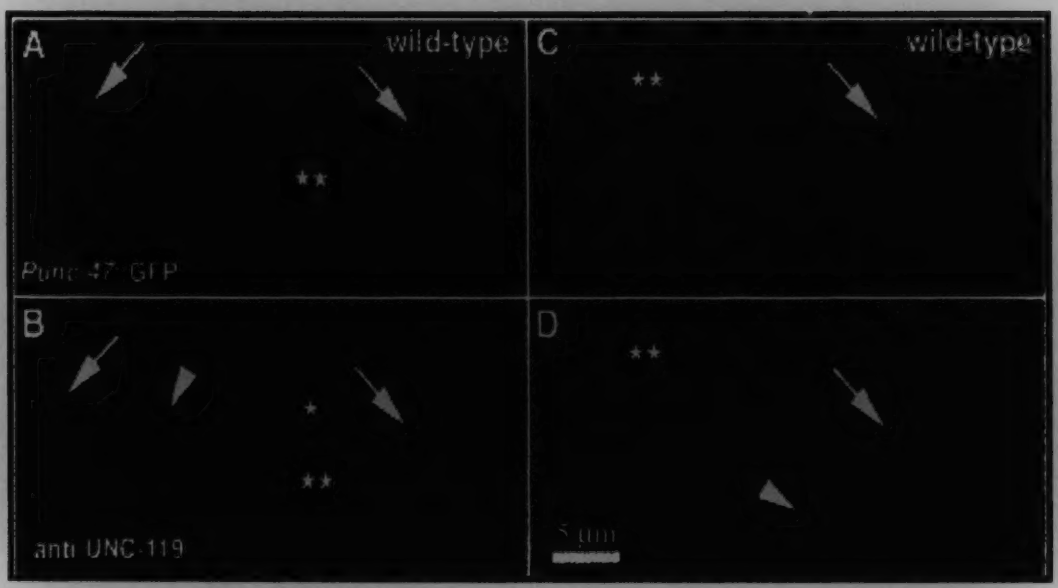

A
wild-type
Punc-47 GFP
B
anti UNC-119
C
wild-type
D
5 µm

Figure 3. A phylogenetic tree of the UNC-119 protein family. The tree of UNC-119 proteins (identified prior to this project) was made using CLUSTAL X (Thompson *et al.* 1997) and the neighbor-joining method (Saitou and Nei 1987). The relative branch lengths of the tree reflect the evolutionary distances between the homologues.

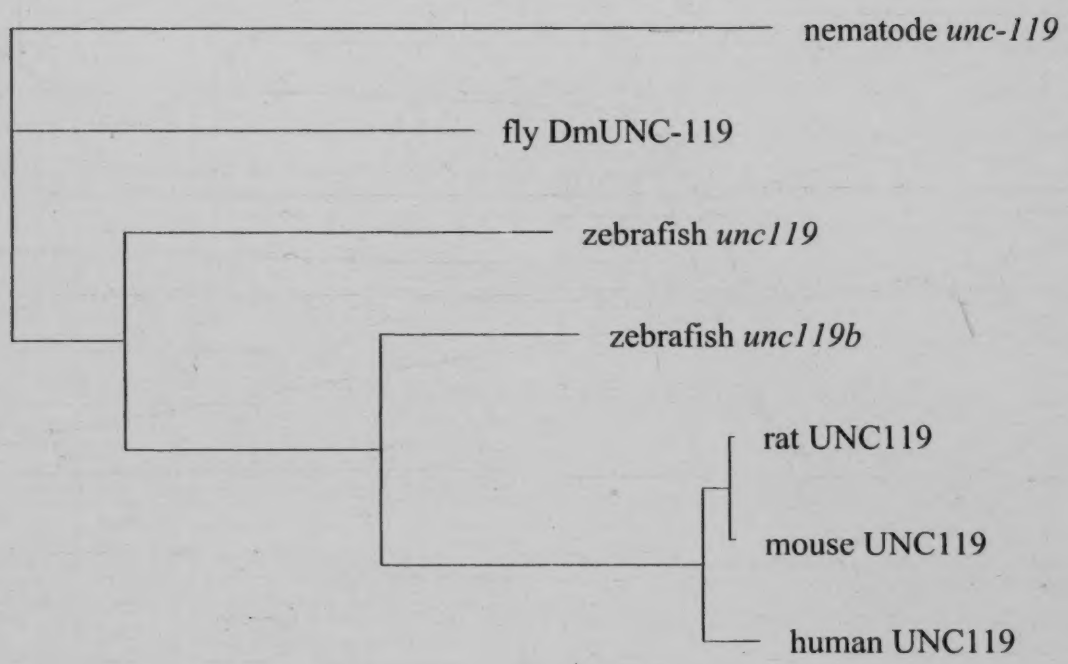

nematode *unc-119*
fly DmUNC-119
zebrafish *unc119*
zebrafish *unc119b*
rat UNC119
mouse UNC119
human UNC119

Figure 4. A comparison of the amino acid sequences of UNC-119 proteins in different organisms. The sequences were aligned using CLUSTAL X with the default settings (Thompson *et al.* 1997). Residues identical or similar among all sequences are highlighted in black and gray, respectively. Close inspection reveals an even more striking conservation, because in many positions (arrows) only a single family member has an amino acid substitution that prevents it from being shaded in that colour.

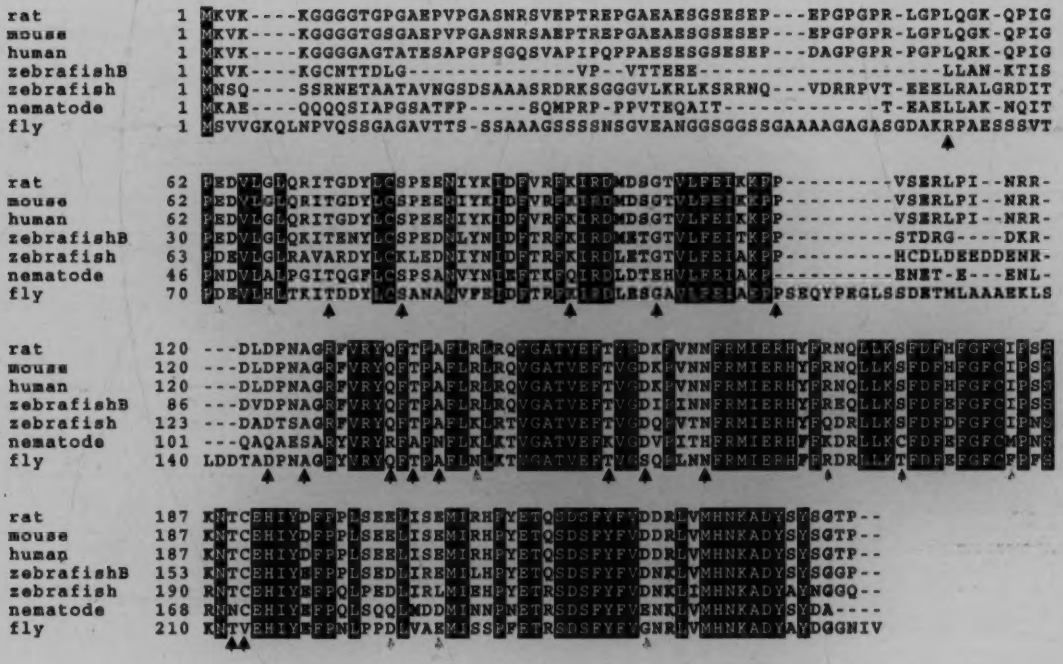

rat
mouse
human
zebrafishB
zebrafish
nematode
fly

Figure 5. DmUNC-119 expression during embryogenesis in *Drosophila melanogaster.* A) Ventral view of a late stage 11 embryo. Anterior is to the left. Transcript accumulation is first detected uniformly in the neuromeres of the central nervous system at this stage. Uniform and ubiquitous expression is maintained throughout the nervous system for the remainder of embryogenesis. B) Sagittal view (anterior to the left, dorsal to the top) of a stage 16 embryo. C) Ventral view (anterior to the left) at stage 13. VNS= ventral nervous system. From Maduro *et al.* 2000, figure 3.

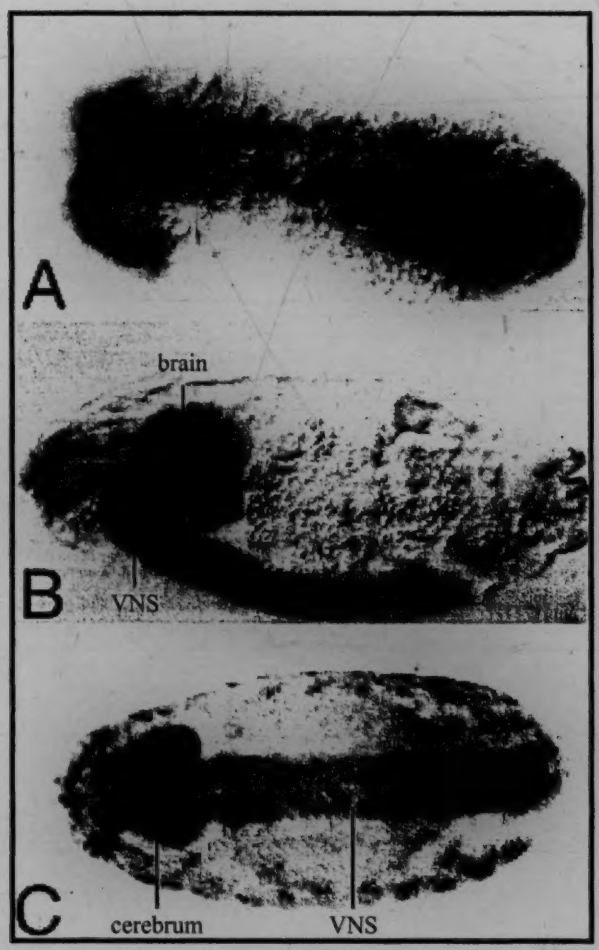

A
brain
B
VNS
C
cerebrum
VNS

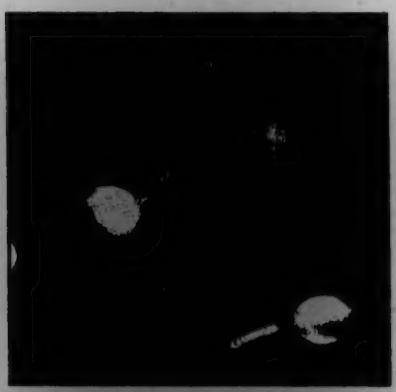
Figure 6. *Danio rerio unc119* and *unc119b* expression and protein localization. *Unc119* and *unc119b* are ubiquitously expressed at the transcriptional level despite distinct protein localization in embryos. A) *Unc119 in situ* hybridization of early (4 hpf) embryos reveals ubiquitous expression. α-Unc119 immuno-fluorescent staining reveals specific localization to the ganglion cell layer (GCL) of the larval (96 hpf) differentiated retina (B) and to several neural structures in the body trunk (C) throughout development. D) *Unc119b in situ* hybridization of early (4 hpf) embryos also reveals ubiquitous expression. α-Unc119b immuno-fluorescent staining reveals specific localization to the inner plexiform layer (IPL) of the larval differentiated retina (E) and to several slow twitch muscles of the body (F) throughout development. hpf= hours post fertilization. Courtesy of A. Manning.

Unc119

Unc119b

A.

D.



B.

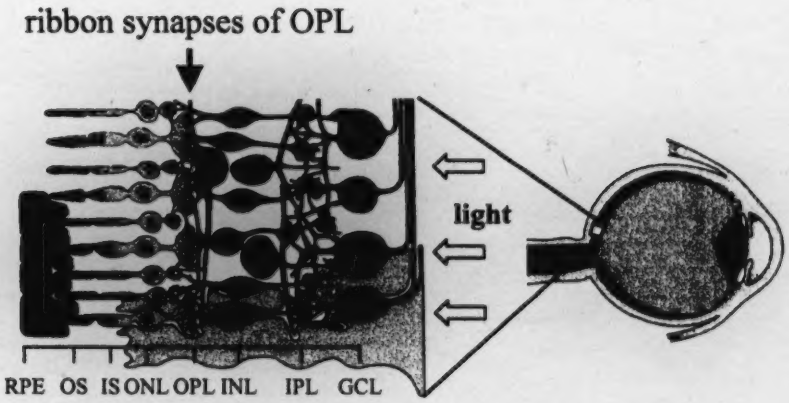
E.

C.

F.

Figure 7. Mammalian UNC119 and ARL2 co-localize in the adult retina. A) The human eye. The different layers of the retina are indicated, including the ribbon synapse, the site of UNC119 enrichment. B) UNC119 and ARL2 co-localize in the OPL of the retina and in the IS of the photoreceptors. Double immuno-fluorescence of a 5 month old rat retina with ARL2 (green) and UNC119 (red). i) A phase micrograph of the retinal layers. ii) α-ARL2 immuno-fluorescent staining demonstrates that ARL2 is present predominantly in the GCL, INL, OPL, and IS, and to a lesser extent in the IPL. iii) α-UNC119 immuno-fluorescent staining demonstrates that UNC119 is present in the OPL and IS layers. GCL= ganglion cell layer; INL= inner nuclear layer; IPL= inner plexiform layer; IS= inner segment; ONL= outer nuclear layer; OPL= outer plexiform layer; OS= outer segment; RPE= retinal pigment epithelium. Figure 7B adapted from Kobayashi *et al.* 2003, figure 3.

A.


ribbon synapses of OPL
light
RPE OS IS ONL OPL INL IPL GCL

B.

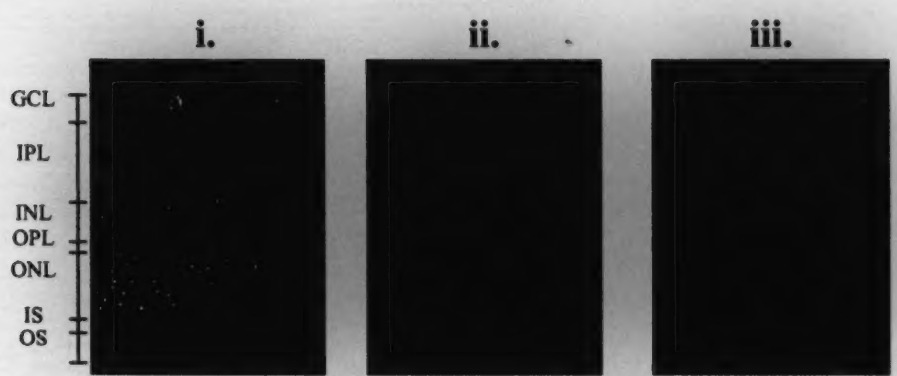

i.
ii.
iii.
GCL
IPL
INL
OPL
ONL
IS
OS

Figure 8. Amino acids that may mediate the UNC119-ARL2 interaction are conserved across UNC-119 proteins. Residues of UNC119 that are predicted to be important for the ARL2 interaction based on the crystal structure of ARL2 with PDEδ (the distant UNC119 relative) are indicated by arrowheads. Arrowheads shaded in black (identical) or in gray (similar) are conserved between UNC119 and PDEδ. The overall similarity between UNC119 and PDEδ for these residues is 10/13, or 77%. As demonstrated above the UNC-119 family alignment (shaded as in figure 4), 9 of these residues are perfectly conserved across the UNC-119 protein family (*), 1 is similar (◊), and the other 3 are conserved across vertebrates (√).

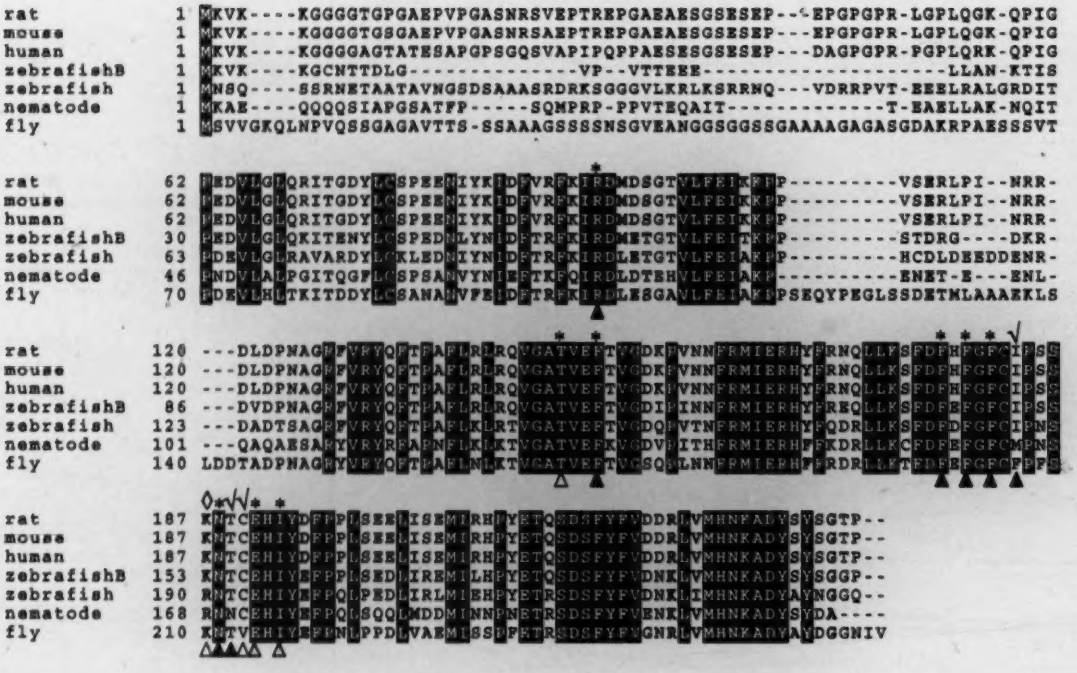
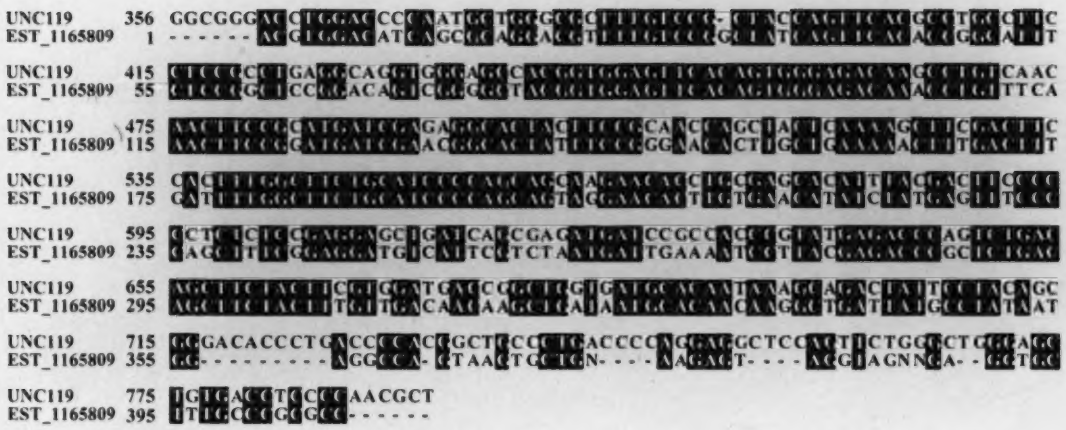

rat
mouse
human
zebrafishB
zebrafish
nematode
fly

Figure 9. Human UNC119B: A homologue of UNC119. A) A nucleotide alignment of an UNC119-like EST (gi:1165809) to the corresponding region (positions 356-792) of the UNC119 mRNA transcript (gi:4103879). They are 71% identical. Nucleotides were aligned using CLUSTAL W (1.74) (Thompson *et al.* 1994) with the default settings. Black text denotes identical residues. B) An amino acid comparison of human UNC119 and the predicted UNC119B protein based on the ORF present in a sequence assembly of overlapping ESTs. Both are 56% identical. The sequences were aligned using CLUSTAL W (1.74) with the default settings (Thompson *et al.* 1994). Residues identical or similar between both sequences are highlighted in black and gray, respectively.
ORF= open reading frame.

A.



B.

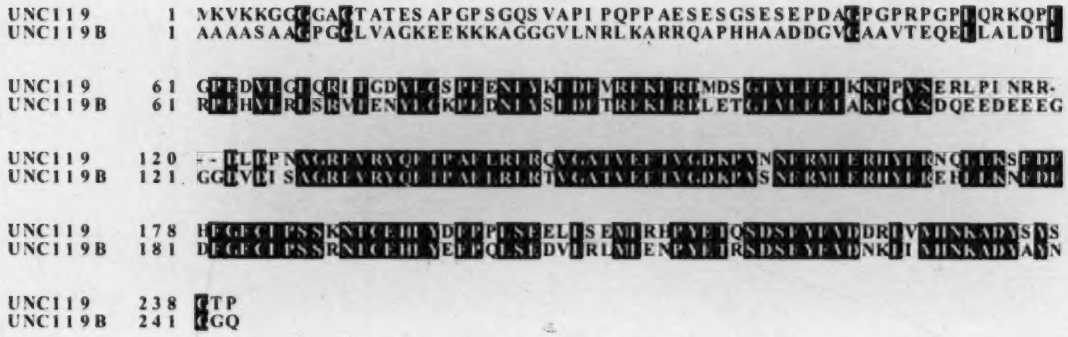
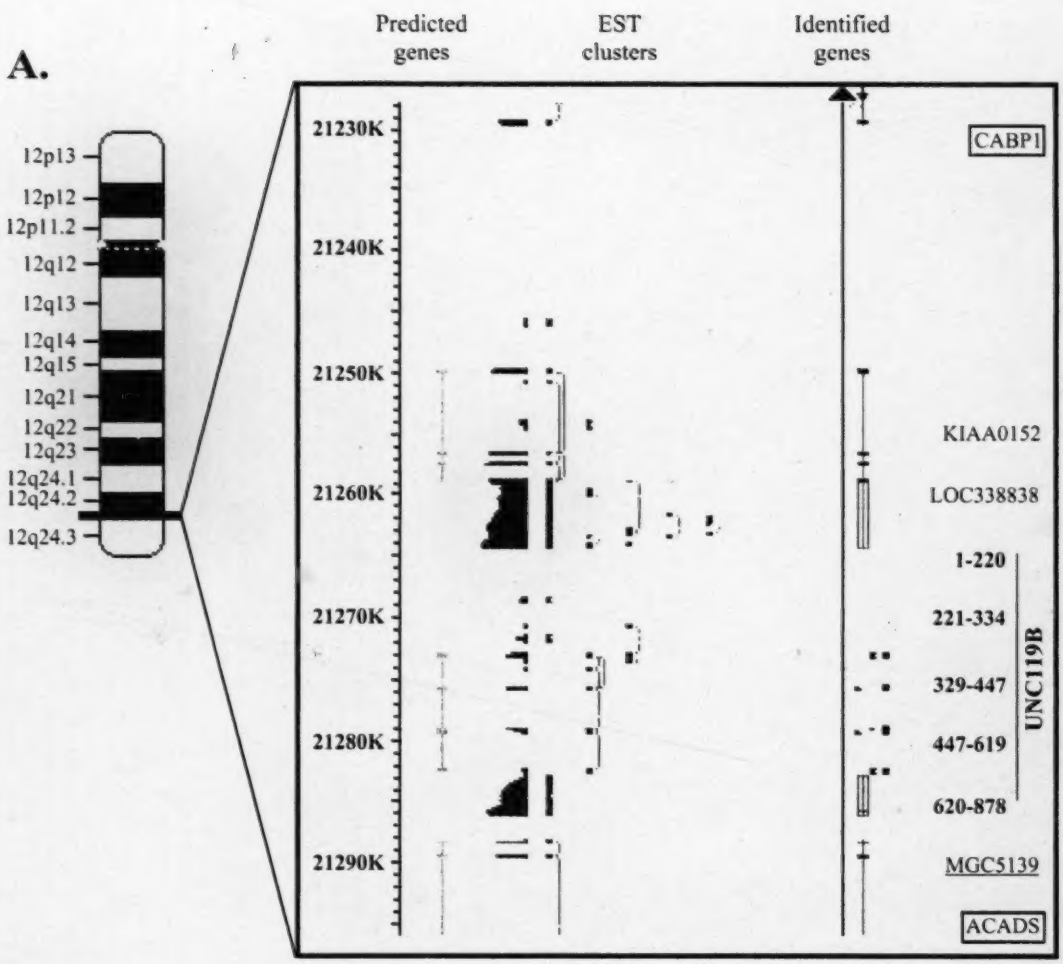


Figure 10. The structure of human UNC119B. A) The UNC119B coding sequence (878 nucleotides) maps to human chromosome 12q24.31 and contains at least 5 exons. In the close-up on the right, the locations of the UNC119B coding DNA are shown in bold, with position 878 representing the codon for the amino acid at the carboxyl end of the protein. The position of the adjacent predicted protein MGC5139 is underlined. The tightly linked genes flanking UNC119B that are conserved between human and mouse are boxed. In the column representing the locations of identified clusters of ESTs, the thickness of the lines is a representation of the number of similar ESTs mapping to that specific site. Figure adapted from the human genome database at the URL (http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?org=hum&MAPS=blast&db=genome%20&na=1&gnl=ref%7CNT_009775.12%7CHs12_9932&gi=29744945&RID=1050710982-020706-12828&segs=11593997-11594256,11584576-11584795,11590947-11591119,11590699-11590817,11587351-11587464&seal=5584°C 6D62C9F48080D5657D727CBA0D7). B) The structure of UNC119B on chromosome 12q24.31 demonstrating the relative distances between exons. The arbitrary numbering on the scale indicates nucleotide distances along the chromosome. Thick horizontal lines represent exons.


A.
Predicted genes
EST clusters
Identified genes
12p13
12p12
12p11.2
12q12
12q13
12q14
12q15
12q21
12q22
12q23
12q24.1
12q24.2
12q24.3
21230K
21240K
21250K
21260K
21270K
21280K
21290K
CABP1
KIAA0152
LOC338838
1-220
221-334
329-447
447-619
620-878
UNC119B
MGC5139
ACADS

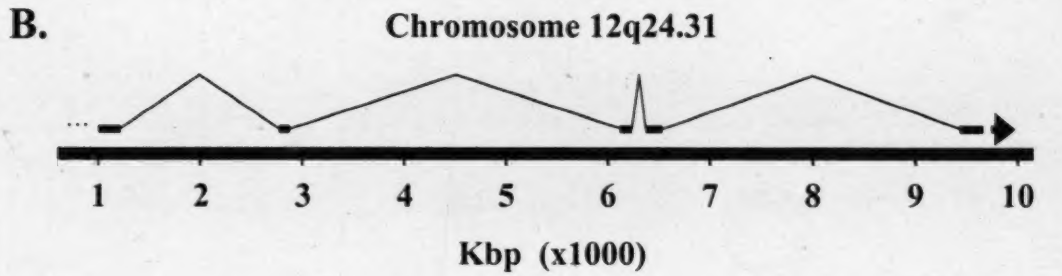

B.
Chromosome 12q24.31
1
2
3
4
5
6
7
8
9
10
Kbp (x1000)

Figure 11. Mammalian UNC119B. A) The predicted 5' end of the human UNC119B protein. The single letter code for the first predicted 8 amino acids of UNC119B based on the genomic sequence of human chromosome 12 are shown. They are not represented in the translation of the assembled UNC119B sequence from ESTs and cDNA clones. B) An amino acid alignment of mammalian UNC119B proteins. Sequences presented were obtained from genomic databases and direct sequencing. The mouse sequence is based on a translation of mouse neonate cerebellum cDNA clone (gi:26349818). The rat protein (gi:27666035) was predicted by automated analysis based on the NCBI chromosome 12 supercontig. The human sequence is based on a translation of an alignment of multiple ESTs that has been confirmed by direct sequencing of the entire clone of the EST (gi:1165809) with the addition of the predicted 5'end shown in A. The sequences were aligned using CLUSTAL W (1.74) with the default settings (Thompson *et al.* 1994). Residues identical or similar among at least 2 of the sequences are highlighted in black and gray, respectively. All 3 homologues are 89% identical. Each rodent protein is 90% identical to human UNC119B, and they are 97% identical to each other.

A. **"MSGSNPKA"**

B.

mouse 1 MSGSNPKAATAGSQAGPGGLVAGKEEKKKAGGGVLNRLKARRQGPPHTPDDGSGA
rat 1 MSGSNPKATTAGSQAGPGGLVAGKEEKKKAGGGVLNRLKARRQGPPHTSDDGPGT
human 1 MSGSNPKAVAAASAAGPGGLVAGKEEKKKAGGGVLNRLKARRQAPFLAADDGVGA
mouse 56 AVTEQELLALDTIRPEHVLRLNRVTENYLCKPEDNVYSIDFTRFKIRDLETGTVI
rat 56 AVTEQELLALDTIRPEHVLRLNRVTENYLCKPEDNVYSIDFTRFKIRDLETGTVI
human 56 AVTEQELLALDTIRPEHVLRLSRVTENYLCKPEDNTYSIDFTRFKIRDLETGTVI
mouse 111 FEIAKPCISDQDQDAEEESVDVDISVGRFVRYQFTPAFLRLRTVGATVEFTVGDR
rat 111 FEIAKPCISDQDQDTEEESVDVDISVGRFVRYQFTPAFLRLRTVGATVEFTVGDR
human 111 FEIAKPCVSDQEEDEEEGGGIDVDISAGRFVRYQFTPAFLRLRTVGATVEFTVGDK
mouse 166 PVTGFRMIERHVFRERLLKTFDFDFGFCTPSSRNTCEIIYEFPQLSFDVTRLMIE
rat 166 PVTGFRMIERHVFRERLLKTFDFDFGFCTPSSRNTCEIIYEFPQLSFDVTRLMIE
human 166 PASNFRMIERHVFRFHLLKNFDFDFGFCTPSSRNTCEIIYEFPQLSFDVTRLMIE
mouse 221 NPYETRSDSFYFVDNKLVMHNKADYAYNGGQ
rat 221 NPYETRSDSFYFVDNKLIMHNKADYAYNGGQ
human 221 NPYETRSDSFYFVDNKLIVHNKADYAYNGGQ

Figure 12. The genomic location of mouse UNC119B. The full-length mouse UNC119B transcript (gi:26349818) maps to mouse chromosome 5 at cytoband E5. In the close-up on the right, the locations of the UNC119B transcript are shown in bold with the exception of the suspected coding exon that does not map to any region of the April 2003 version of the genome. The position of the tightly linked adjacent genes that demonstrate that this region shows conserved synteny with human chromosome 12 are boxed, and the orientation of the conserved region is reversed with respect to figure 10. LOC231657 represents a short protein predicted by automated processing based on 4 ESTs. In the column representing the locations of identified clusters of ESTs, the thickness of the lines is a representation of the number of similar ESTs mapping to that specific site. Figure adapted from the mouse genome database at the URL (http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?org=mouse&MAPS=blast&db=genome%20&na=1&gnl=ref%7CNT_039312.1%7CMm5_39352_30&gi=28548132&RID=1050800013-016737-23563&segs=1649185-1652120,1661046-1661354,1657031-1657144,1653867-1653980&seal=F9B7238EE331F3947AFD13B444AFFF44).

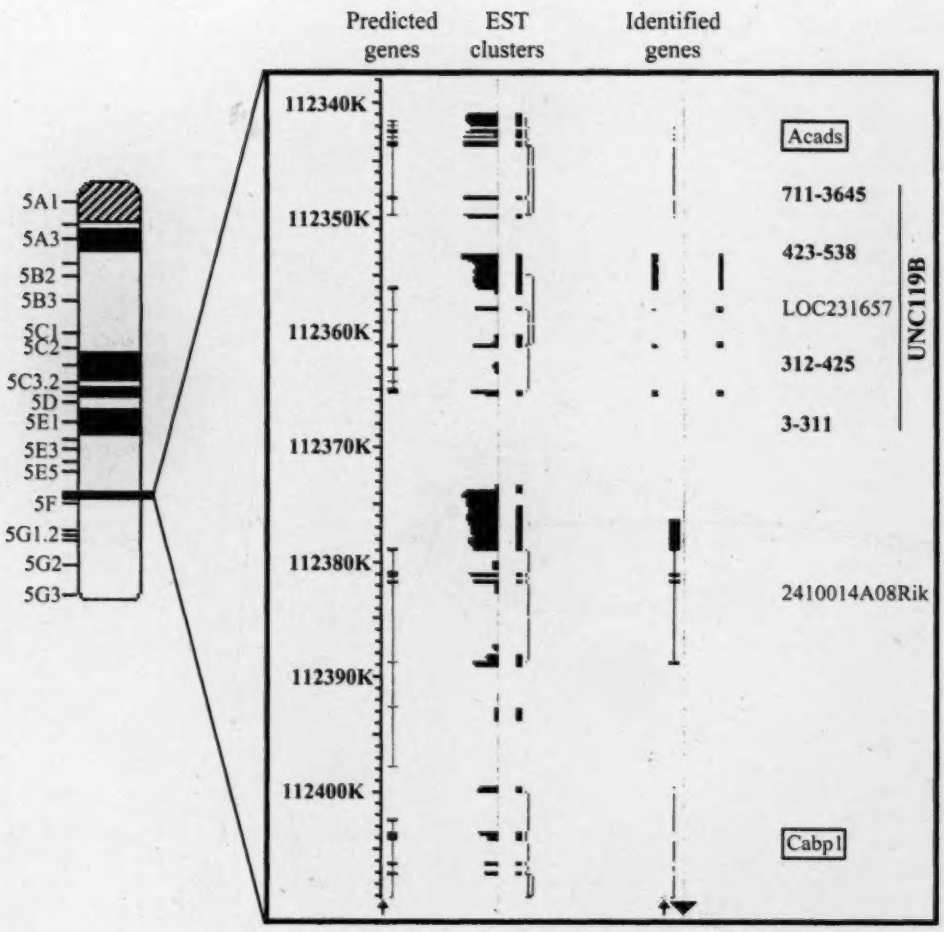
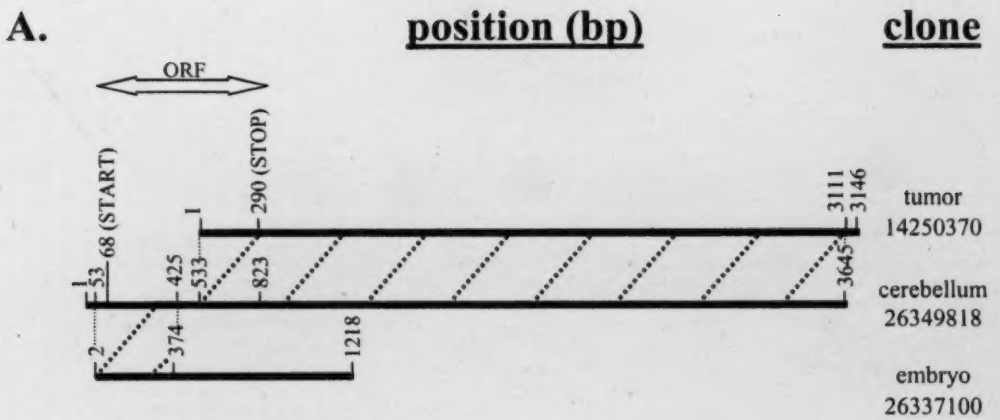

Predicted genes
EST clusters
Identified genes
5A1
5A3
5B2
5B3
5C1
5C2
5C3.2
5D
5E1
5E3
5E5
5F
5G1.2
5G2
5G3
112340K
112350K
112360K
112370K
112380K
112390K
112400K
Acads
711-3645
423-538
LOC231657
312-425
3-311
UNC119B
2410014A08Rik
Cabp1

Figure 13. The gene structure of mouse UNC119B. A) Incomplete overlap of different mouse UNC119B cDNA clones suggests the possibility of alternate transcripts. Diagonal hatched lines are drawn between overlapping sequences from each clone. Clone 14250370 is derived from mammary tumor. Clone 26349818 was obtained from a neonate cerebellum library. Clone 26337100 was derived from RNA extracted from 9.5 day embryo parthenogenote tissue. B) The structure of the mouse UNC119B clones on chromosome 5E5 demonstrating the relative distances between exons. The numbering on the scale is arbitrary. The reference number (gi) for each clone is indicated in bold. The un-mapped portion of the putative fourth exon of the full-length clone (gi:26349818) and the 5' end of the partial clone (gi:14250370) is arbitrarily positioned and shown in brackets. The thick horizontal lines represent exons, and the ORFs of each clone (position 1-290 of 14250370, 68-823 of 26349818 and 17-670 of 26337100) are demonstrated by an overlying second thin line. The ORF of 26337100 is different from the one present in the other clones after postions 393. A suspected facultative splice donor site is indicated by the arrow. The 3' end of the UNC119B transcript is to the left; 5' is to the right. ORF= open reading frame.


A.
position (bp)
clone
ORF
68 (START)
290 (STOP)
1
53
425
533
823
3111
3146
3645
tumor
14250370
cerebellum
26349818
2
374
1218
embryo
26337100

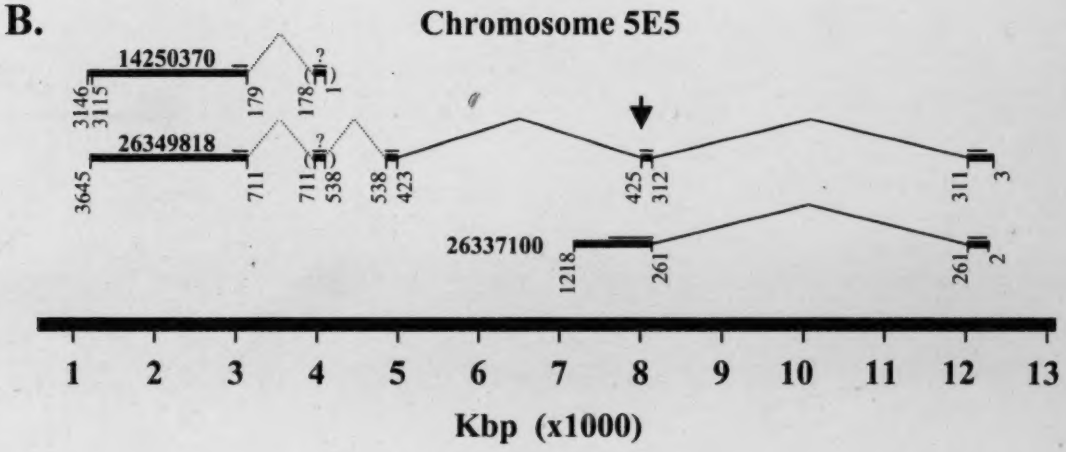

B.
Chromosome 5E5
14250370
3146
3115
179
178
1
3
26349818
3645
711
711
538
538
423
425
312
311
3
26337100
1218
261
261
2
1
2
3
4
5
6
7
8
9
10
11
12
13
Kbp (x1000)

Figure 14. UNC119 and UNC119B are widely expressed in the human fetus. A human fetal tissue northern blot was probed sequentially with UNC119 (A), UNC119B (B), and ß-actin (C). The blot was stripped between uses. A and B are both aligned against the size standard in kilobases (Kb) shown. Exposure times were 35 hours and 1 week respectively for UNC119 and UNC119B. More stringent washes for UNC119 (0.1X SSC/0.1% SDS at 73°C) do not attenuate or change the ratio of any visible signal, confirming the specificity of the detected products.

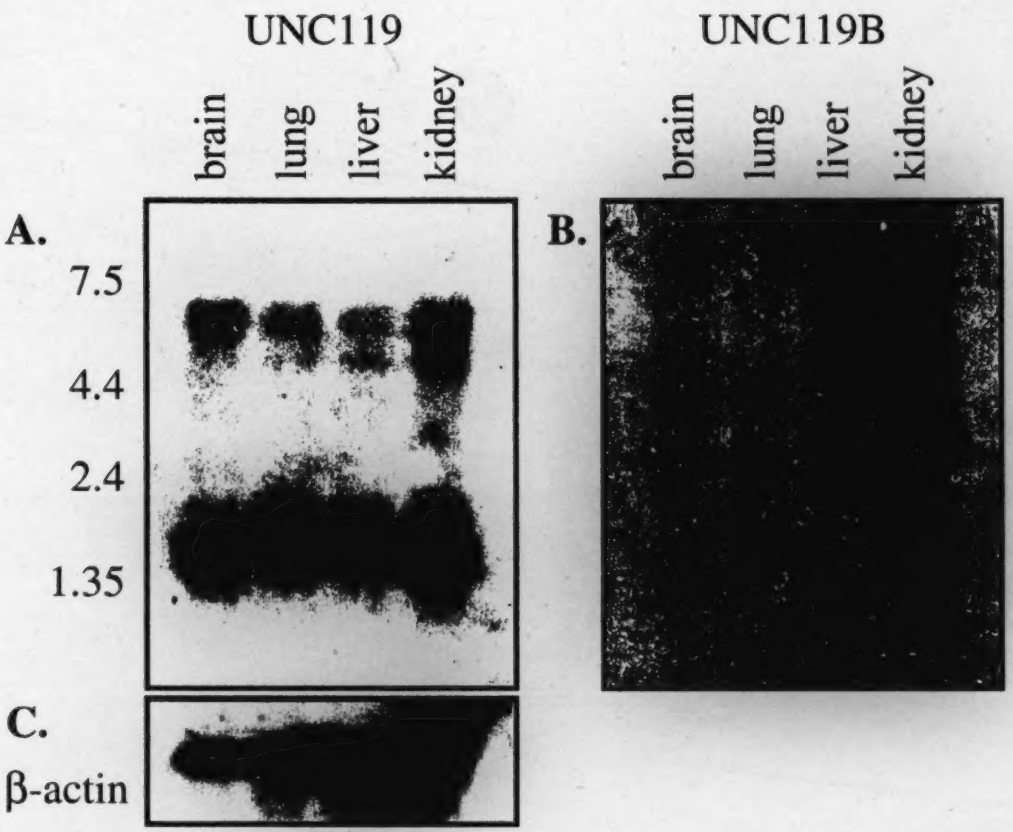

UNC119
UNC119B
brain
lung
liver
kidney
brain
lung
liver
kidney
A.
B.
7.5
4.4
2.4
1.35
C.
β-actin

Figure 15. UNC119 and UNC119B mRNA expression in mouse E10.5 embryos. Embryos were collected 10 days post coitum from the CD-1 breeder strain and probed with UNC119 antisense (A), UNC119 sense (B), UNC119B antisense (C) or UNC119B sense (D) riboprobes. Nucleotide sequences of the fragments from which probes were transcribed are shown in Appendix 3. Embryos were hybridized for 14 hours and colour was developed for 3.5 hours.

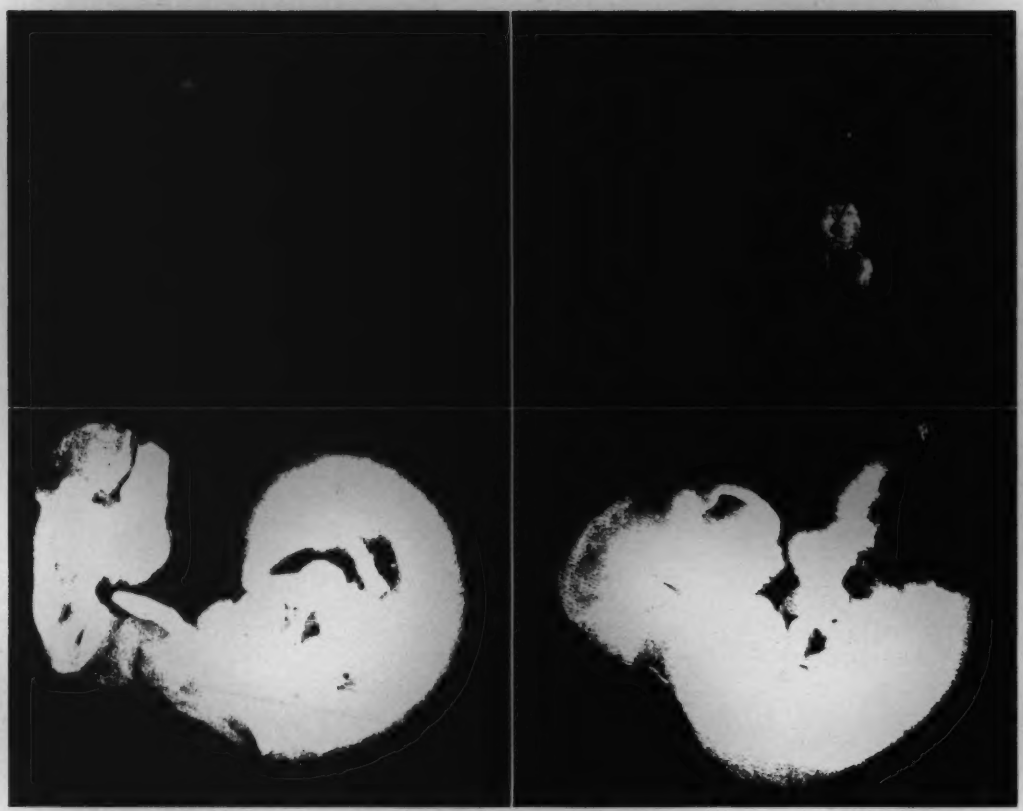


Figure 16. UNC119 and UNC119B mRNA expression in mouse E11.5 embryos. Embryos were collected 11 days post coitum from the CD-1 breeder strain and probed with UNC119 antisense (A), UNC119 sense (B), UNC119B antisense (C) UNC119B sense (D), or *Pax1* (E) riboprobes. *Pax1* is a somite specific positive control. Nucleotide sequences of the fragments from which probes were transcribed are shown in Appendix 3. Embryos were hybridized for 14 hours and colour was developed for 3.5 hours.

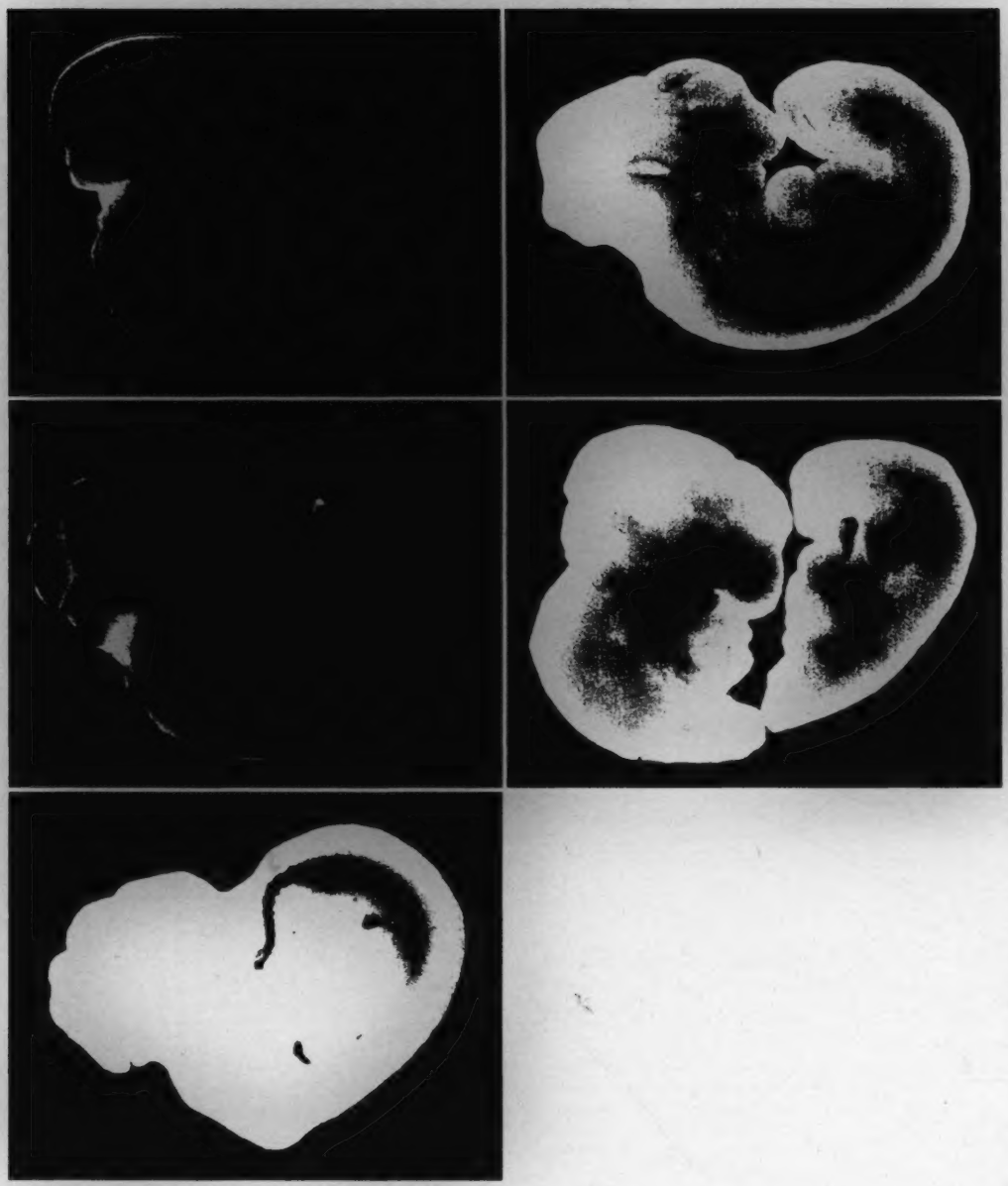


Figure 17. UNC119 and UNC119B mRNA expression in mouse E12.5 embryos. UNC119 expression is more concentrated in the cranium while UNC119B is more prominent in the caudal regions. Embryos were collected 12 days post coitum from the CD-1 breeder strain and probed with UNC119 antisense (A), UNC119 sense (B), UNC119B antisense (C) UNC119B sense (D), or *Pax1* (E) riboprobes. Nucleotide sequences of the fragments from which probes were transcribed are shown in Appendix 3. Embryos were hybridized for 14 hours and colour was developed for 3.5 hours.

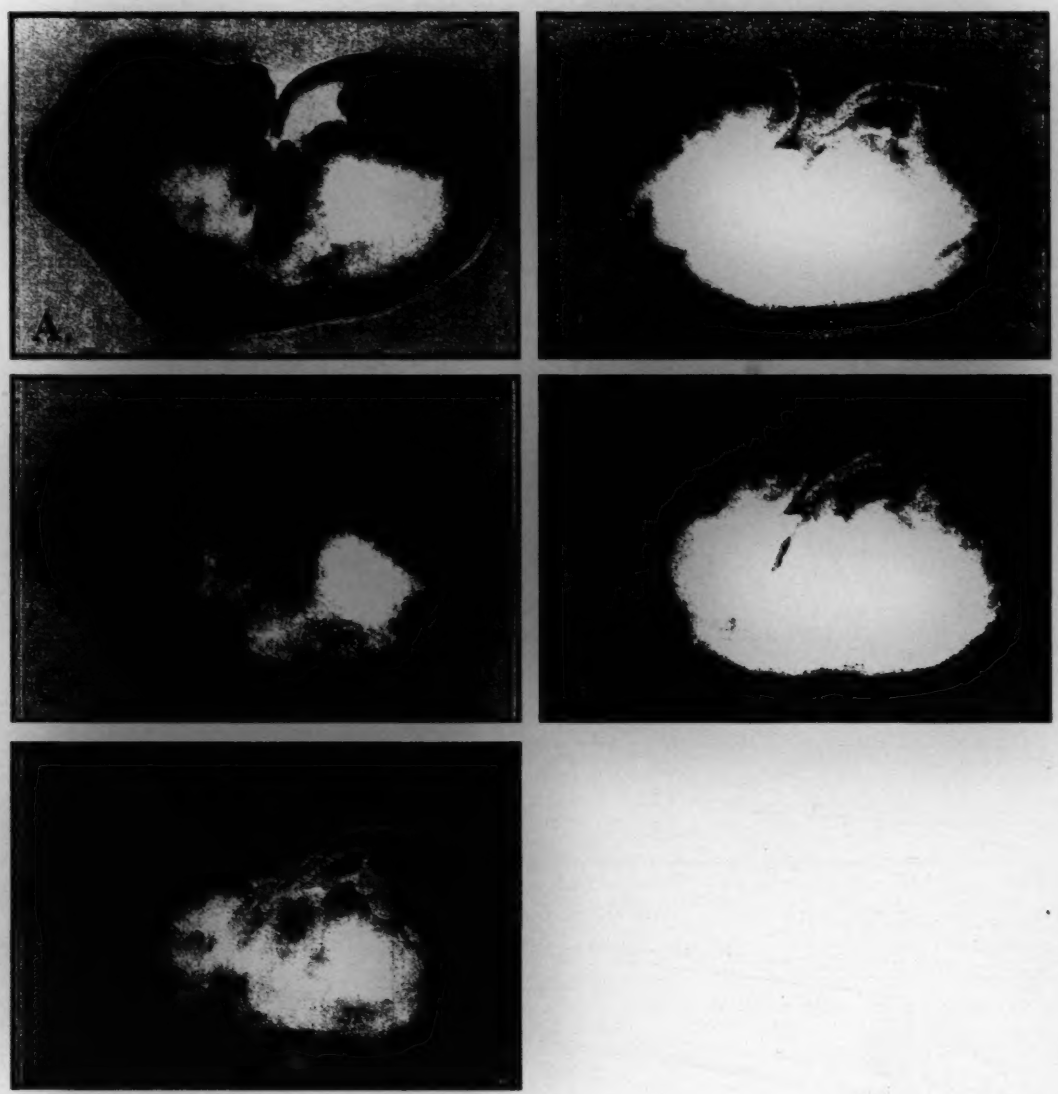

A.

Figure 18. UNC119 *in situ* hybridization of mouse embryonic day 8, 10 and 14 sections. Embryo slide sets (Novagen, Inc.) were prepared using paraformaldehyde fixed, paraffin embedded NIH Swiss strain mouse embryos. They were sectioned at a thickness of 7μm in random orientations from embryos *in utero* at stages E8.5 (A-D) and E10.5 (E-F), and para-sagitally at the same thickness at stage E14.5 (G-H). Nearly adjacent sections were probed with digoxigenin-labeled UNC119 antisense (A, C, E, G) or sense (B, D, F, H) riboprobes. Sections were hybridized at 60°C for 18 hours and developed overnight. Nucleotide sequences of the fragments from which probes were transcribed are shown in Appendix 3. a= liver; b= vertebral bodies; c= shoulder girdle; d= trigeminal ganglion; e= vibrissa; f= mouth; g= intestine. Approximate magnifications are as follows: 17X for A and B; 55X for C and D; 20X for E and F; 7X for G and H.

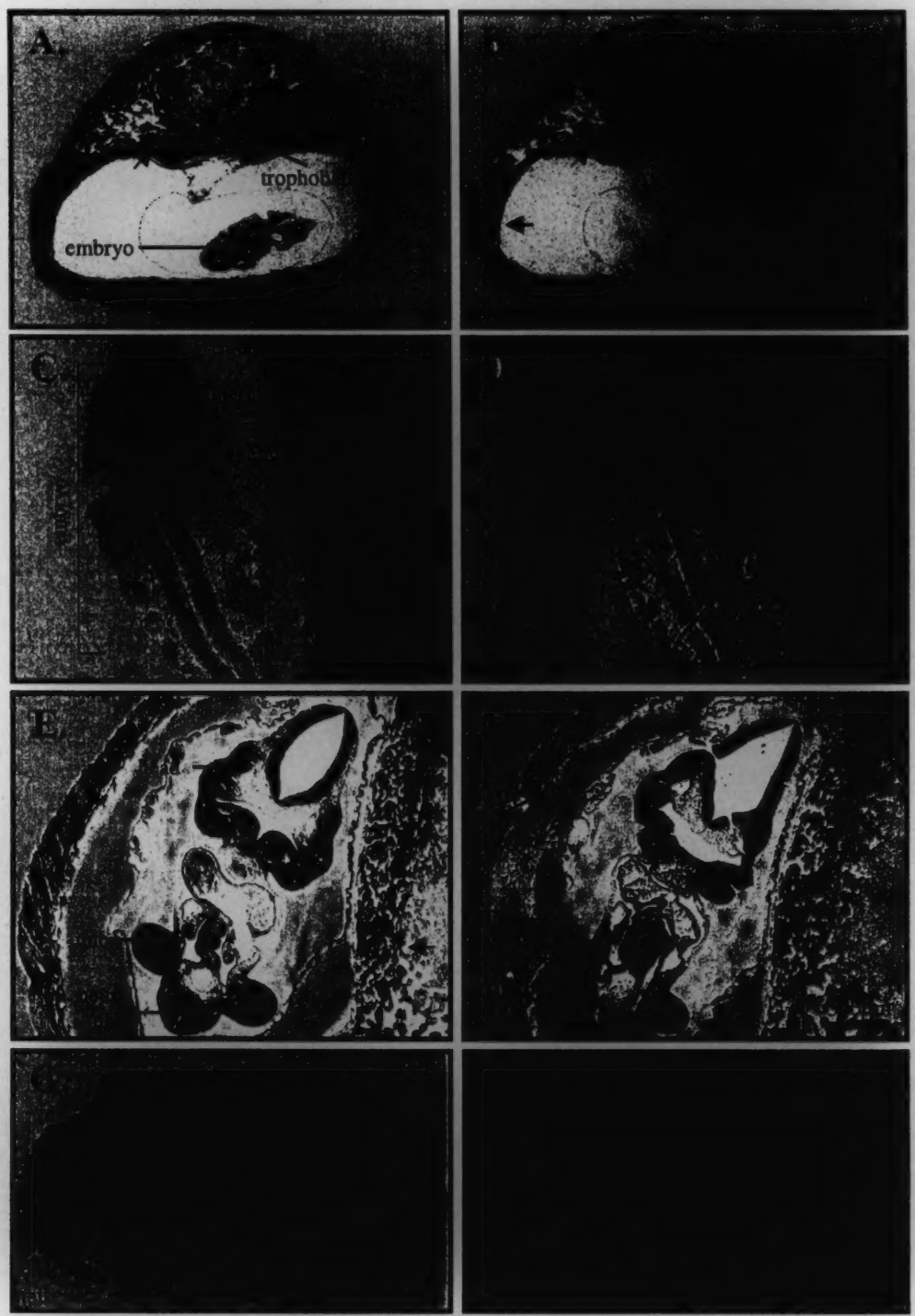

A.
trophob
embryo
C.
E.

Figure 19. UNC119B *in situ* hybridization of mouse embryonic day 8, 10 and 14 sections. Embryo slide sets (Novagen, Inc.) were prepared using paraformaldehyde fixed, paraffin embedded NIH Swiss strain mouse embryos. They were sectioned at a thickness of 7μm in random orientations from embryos *in utero* at stages E8.5 (A-B) and E10.5 (C-D), and para-sagitally at the same thickness at stage E14.5 (E-F). Nearly adjacent sections were probed with digoxigenin-labeled UNC119B antisense (A, C, E) or sense (B, D, F) riboprobes. Sections were hybridized at 50°C for 15 hours and developed overnight. Nucleotide sequences of the fragments from which probes were transcribed are shown in Appendix 3. a= liver; b= vertebral bodies; c= shoulder girdle; d= trigeminal ganglion; e= nose; f= limbs; g= intestine. Approximate magnifications are as follows: 55X for A and B; 17X for C and D; 7X for E and F.

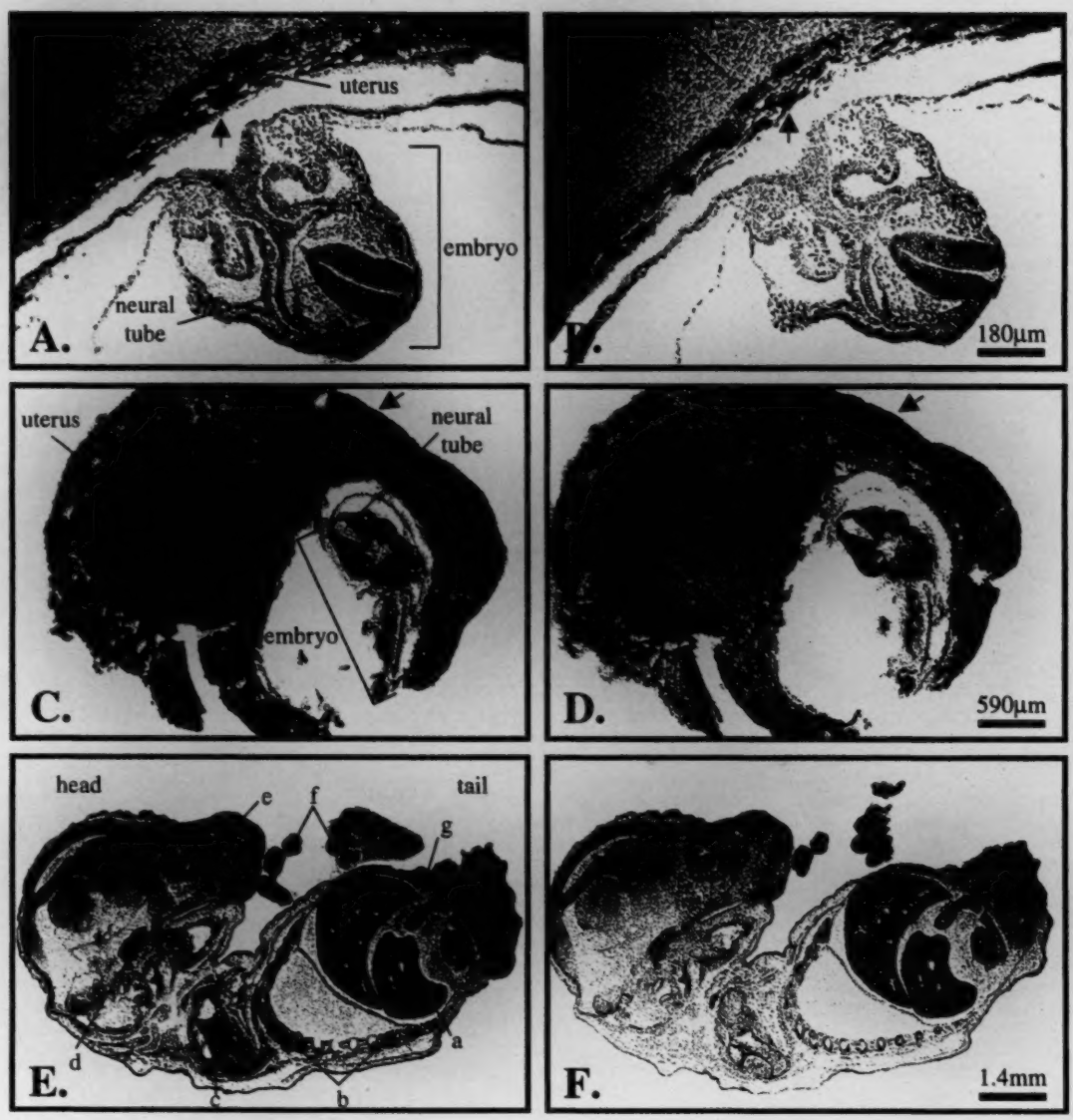
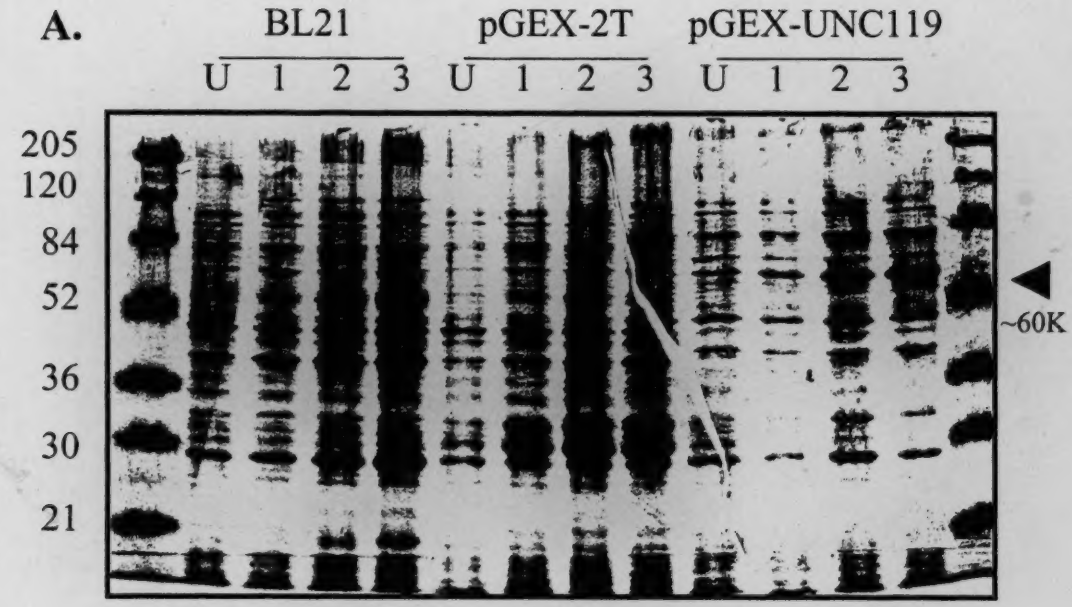

uterus
embryo
neural
tube
A.
180µm
uterus
neural
tube
embryo
C.
D.
590µm
head
tail
e
f
g
a
d
c
b
E.
F.
1.4mm

Figure 20. Induction of UNC119-GST and UNC119B-GST fusion proteins in bacterial cells. A) UNC119-GST migrates at ~60K and begins to be produced 1 hour following induction of the BL21(DE3) bacterial host cells. Production plateaus after 2 hours. B) UNC119B-GST migrates at ~70K and is highly induced after 1 hour. Arrows indicate the induced proteins. BL21 corresponds to untransformed cells; pGEX-2T corresponds to BL21(DE3) cells transformed with an empty GST expression vector; pGEX-UNC119 corresponds to UNC119-GST transformed in BL21(DE3) cells; pGEX-UNC119B corresponds to UNC119B-GST transformed in BL21(DE3) cells; U= uninduced cell lysate; 1, 2, 3= hours post induction.


A.
BL21
pGEX-2T
pGEX-UNC119
U 1 2 3 U 1 2 3 U 1 2 3
205
120
84
52
36
30
21
~60K

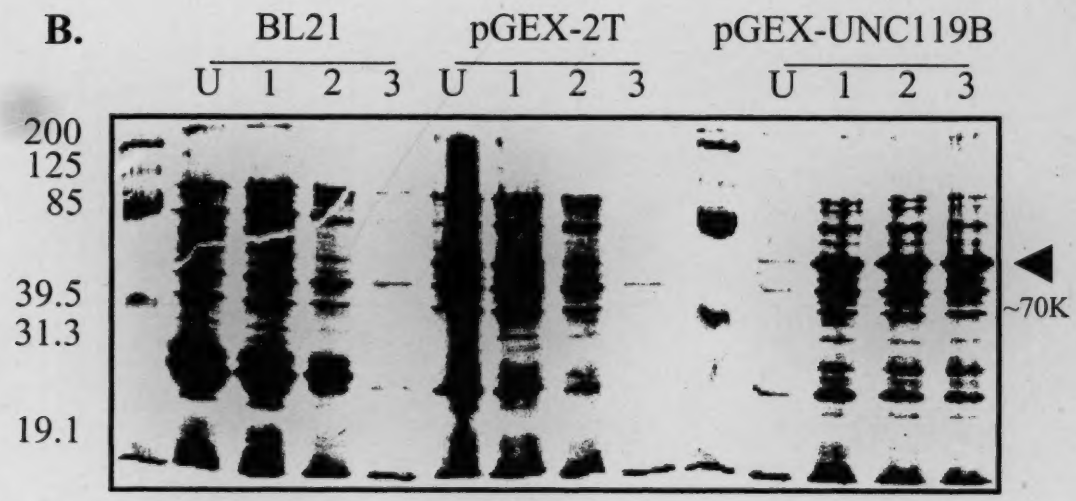
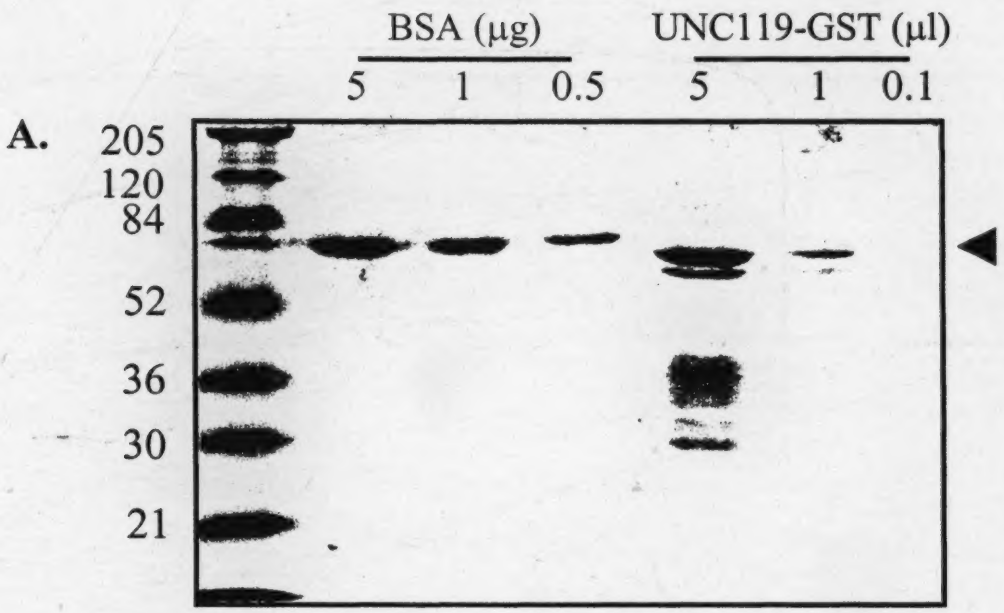

B.
BL21
pGEX-2T
pGEX-UNC119B
U 1 2 3 U 1 2 3 U 1 2 3
200
125
85
39.5
31.3
19.1
~70K

Figure 21. Purification of GST-fusion proteins from the bacterial strain BL21(DE3) following induction. The largest, brightest bands (arrows) migrate at the expected sizes for the full-length fusion proteins. The smaller bands may represent degradation products. A) UNC119-GST. By visual comparison with BSA standard, the UNC119-GST concentration is ~0.2 μg/μl. B) Purified UNC119B-GST. The UNC119B-GST yield is very high and was quantitated to 34μg/μl using spectrometrical analysis. 3μl of the UNC119B-GST protein lysate was fractionated on the gel shown. BSA= bovine serum albumin.


BSA (μg)
UNC119-GST (μl)
5
1
0.5
5
1
0.1
A.
205
120
84
52
36
30
21

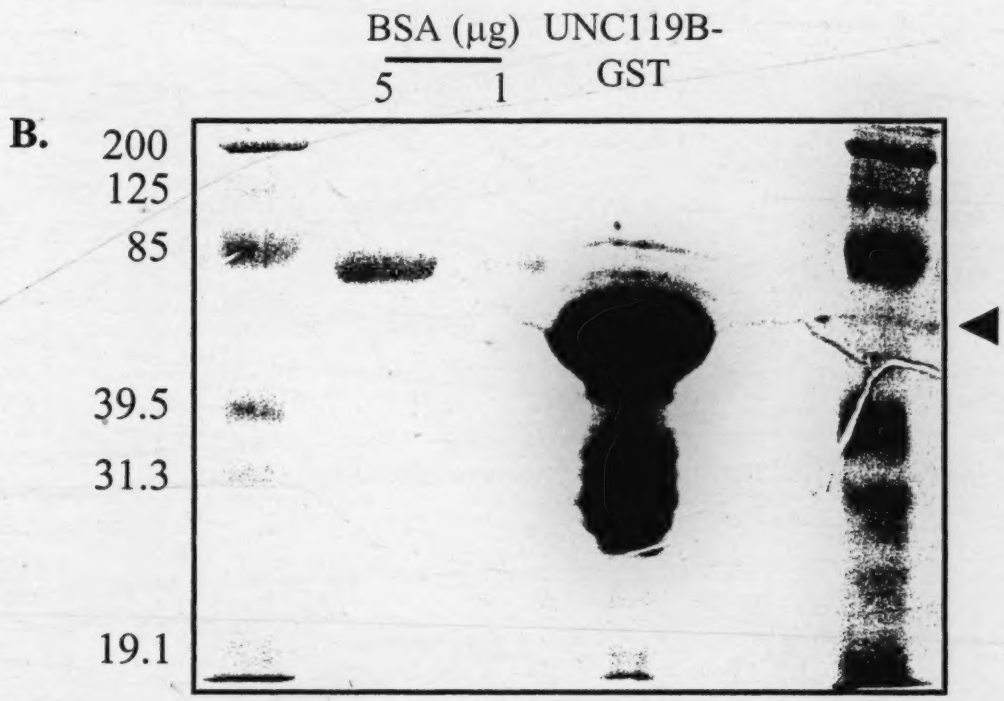
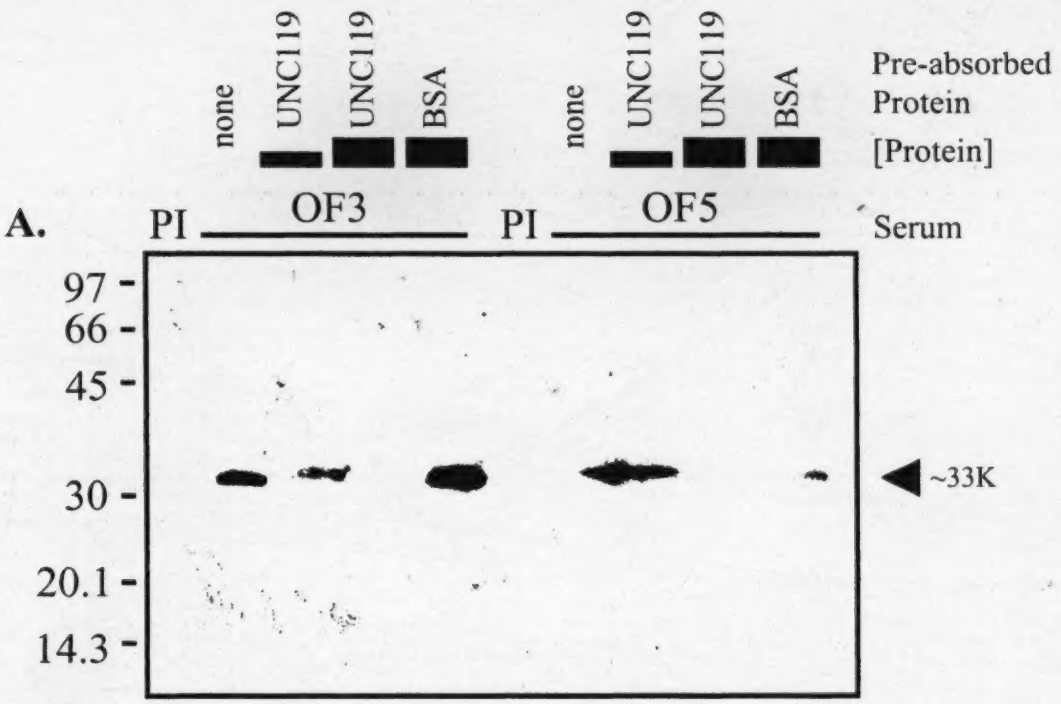

BSA (μg)
UNC119B-GST
5
1
B.
200
125
85
39.5
31.3
19.1

Figure 22. Western blot of α-UNC119 and α-UNC119B antisera against purified recombinant protein. A) 1:10000 dilutions of serum collected from the α-UNC119 injected rabbits OF3 and OF5 recognize 10ng of the UNC119-GST fusion protein used for injection. Serum collected prior to injection (data not shown) detects no signal. B) A 1:500 dilution of serum collected from the α-UNC119B injected rabbits 11-5 and 11-6 strongly recognizes 0.1μg of the UNC119B-GST fusion protein. Even at the high concentrations of sera and protein tested, pre-immune serum (PI) from 11-5 detects no product, and PI serum from 11-6 faintly recognizes trace amounts of a large protein. Only antiserum from rabbit 11-5 has been used for further study. All bands detected are consistent with the purified protein bands demonstrated in figure 21. IS= immune serum; PI= pre-immune serum.

A.
OF3
OF5
107
77
52
35
30
25
B.
11-5
11-6
PI
IS
PI
IS
127
85
40.7
31.9
18.7

Figure 23. Western blot of α-UNC119 antisera against bovine retinal protein. A) A single product of ~33K consistent with previous reports (Higashide *et al.* 1998) is detected by antiserum from each UNC119-GST immunized rabbit (OF3 and OF5). Pre-immune serum (PI) from both rabbits fails to detect any product. Antiserum pre-absorbed with increasing concentrations (7.5×10^{-5} μg/μl and 7.5×10^{-2} μg/μl) of exogenous UNC119-GST or with high concentrations of BSA was also applied in parallel to aliquots of retinal protein to determine the specificity of each antiserum. Signal attenuation is observed only when the antisera are pre-absorbed with UNC119-GST. B) OF3 and OF5 antisera do not cross-react with UNC119B. UNC119B-GST is not capable of attenuating detection of native UNC119 by α-UNC119 antisera when pre-absorbed to either antiserum at similar concentrations (7.1×10^{-5} μg/μl and 7.1×10^{-2} μg/μl). BSA= bovine serum albumin; PI= pre-immune serum.


A.
none
UNC119
UNC119
BSA
none
UNC119
UNC119
BSA
Pre-absorbed Protein [Protein]
PI
OF3
PI
OF5
Serum
97
66
45
30
20.1
14.3
~33K

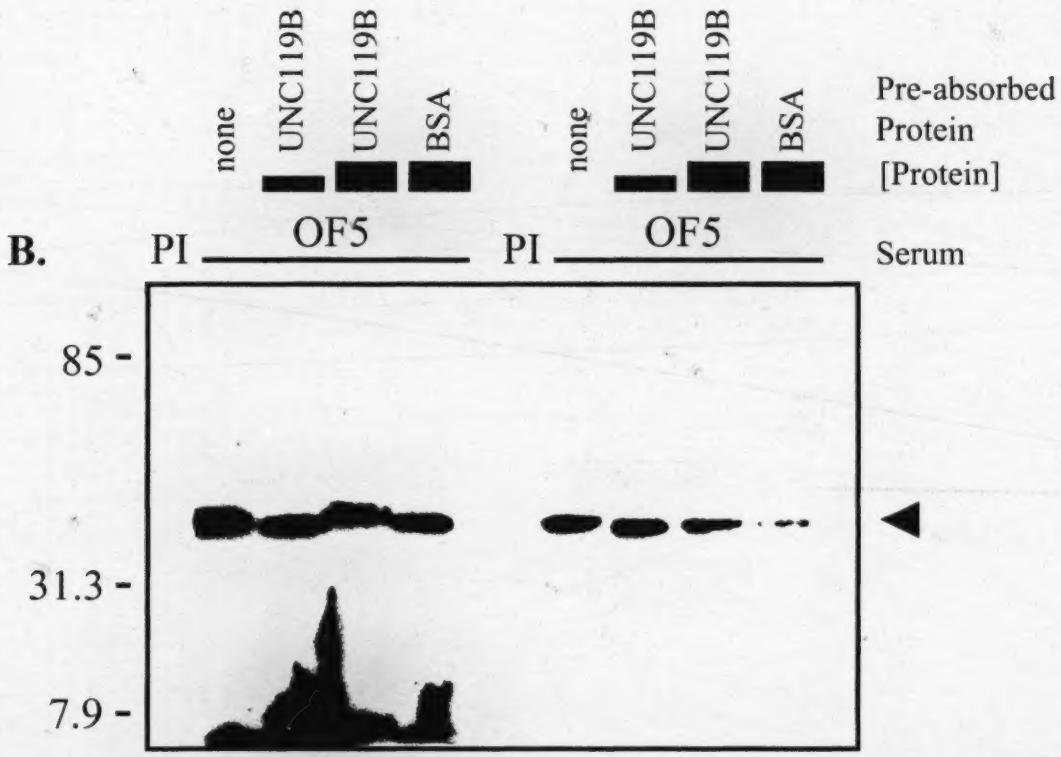

B.
none
UNC119B
UNC119B
BSA
none
UNC119B
UNC119B
BSA
Pre-absorbed Protein [Protein]
PI
OF5
PI
OF5
Serum
85
31.3
7.9

Figure 24. Western blot of purified α-UNC119B against bovine retinal protein. A) Purified 11-5 specifically recognizes a ~55K protein in retinal extracts after a 10 minute exposure. Antiserum was used at a 1:100 dilution, and pre-absorption was performed using 7.5 X 10^{-2} μg protein per μl of antiserum. Signal attenuation is observed only when antiserum is pre-absorbed with UNC119B-GST. Pre-immune serum fails to detect the product. B) Exposure of the same blot for several hours shows faint detection of other proteins by the purified antiserum. 2.25 μl of protein extract was size fractionated per lane. IS= immune serum; PI= pre-immune serum; PS= purified immune serum.

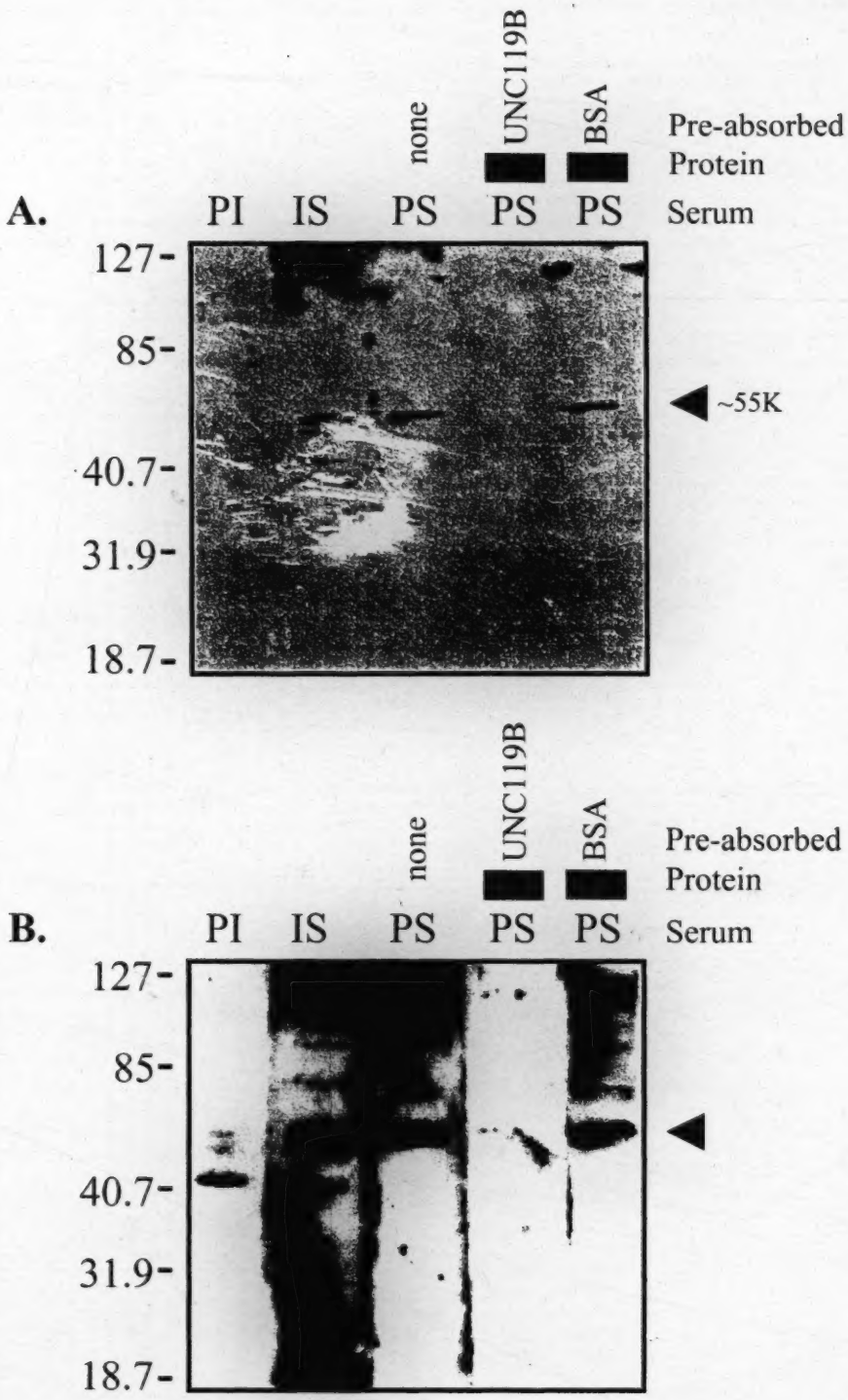
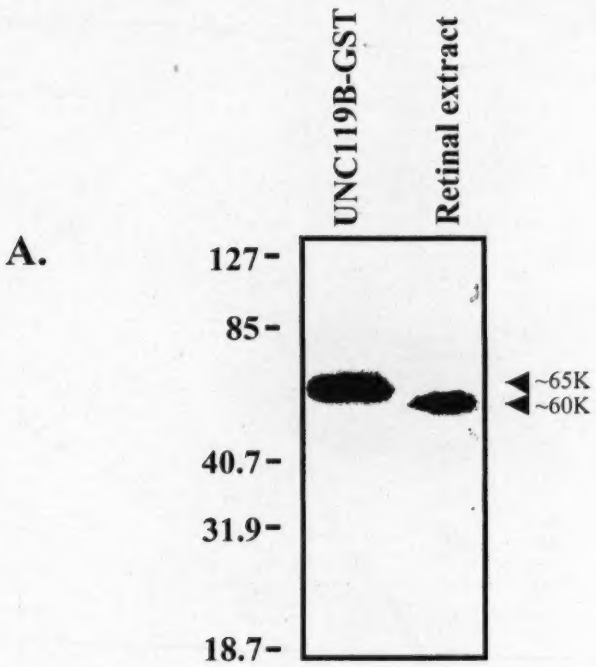

A.
none
UNC119B
BSA
Pre-absorbed Protein
PI
IS
PS
PS
PS
Serum
127-
85-
40.7-
31.9-
18.7-
~55K
B.
none
UNC119B
BSA
Pre-absorbed Protein
PI
IS
PS
PS
PS
Serum
127-
85-
40.7-
31.9-
18.7-

Figure 25. Characterization of purified α-UNC119B. A) The product recognized in bovine retinal protein extracts by purified 11-5 α-UNC119B antiserum migrates more quickly (~60K compared to ~65K) than the GST-fusion form. 1.1 X 10^{-3}µl and 2.25µl of UNC119B-GST and bovine retinal extract were size fractionated respectively in parallel and Western blotted. The blot was separated, and antiserum was used at a 1:1000 dilution against recombinant protein and at a 1:100 dilution against retinal protein. B) α-UNC119 (purified from OF3) and α-UNC119B (purified from 11-5) antisera do not cross-react with the other homologue. Both strongly specifically recognize a single protein consistent with their expected sizes in bovine retinal extracts shown to the left. Upon exposure to saturation, other products appear to be faintly detected by each antiserum, however no evidence of cross-reaction to the other UNC-119 protein is present. Antiserum was applied at a 1:100 dilution, and 2.25µl of protein extract was size fractionated per lane. o/n= overnight exposure (to saturation).


A.
UNC119B-GST
Retinal extract
127
85
40.7
31.9
18.7
~65K
~60K

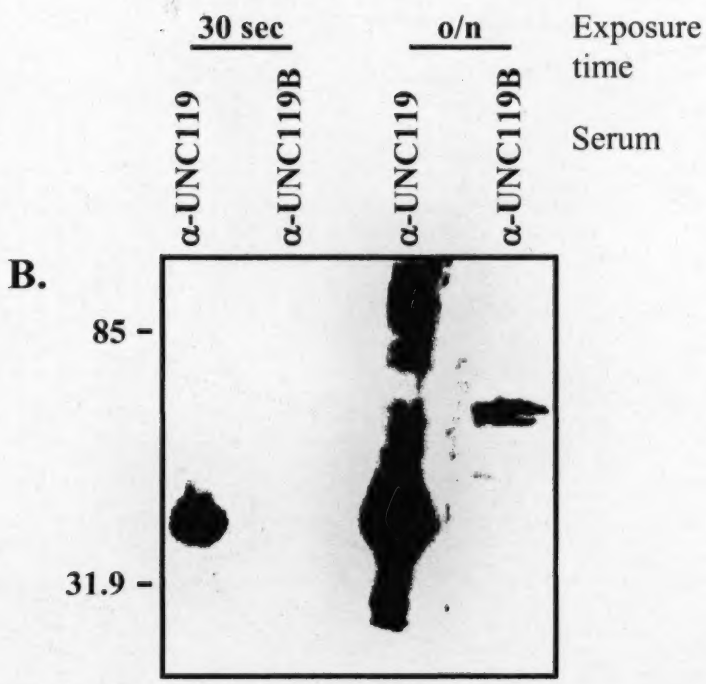

B.
30 sec
o/n
Exposure time
Serum
α-UNC119
α-UNC119B
α-UNC119
α-UNC119B
85
31.9

Figure 26. Western blot of α-UNC119 antisera against adult bovine brain protein. Several products are detected by 1:100 dilutions of each α-UNC119 antiserum following saturation exposures. None of the detected products are specifically out-competed by pre-absorption of either antiserum with increasing amounts of exogenous UNC119-GST. Antiserum was pre-absorbed with increasing concentrations (7.1×10^{-5} or 7.1×10^{-2} μg protein per μl of antiserum) of UNC119-GST or with 7.1×10^{-2} μg/μl BSA. Clear signal attenuation by any pre-absorption is not observed. Pre-immune serum (PI) from both fails to detect the product. Arrowheads indicate the bands detected by each antiserum that migrate at a rate close to the expected size (~33K) for native UNC119. PI= pre-immune serum.

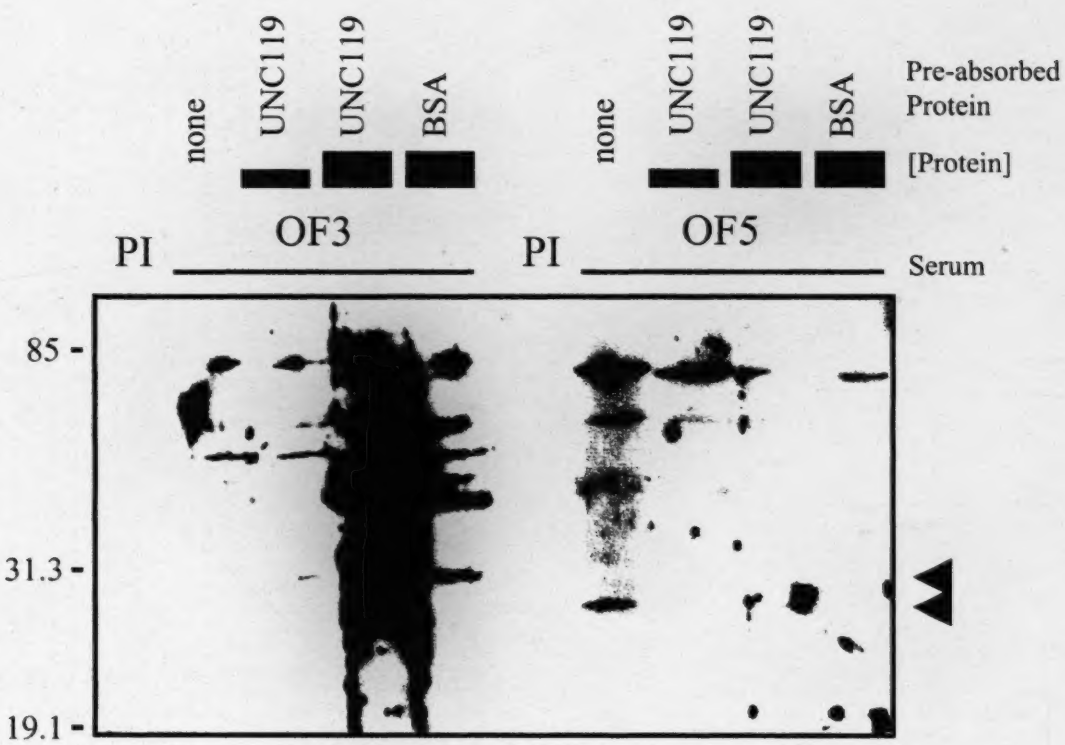

none
UNC119
UNC119
BSA
none
UNC119
UNC119
BSA
Pre-absorbed Protein
[Protein]
PI
OF3
PI
OF5
Serum
85 -
31.3 -
19.1 -

Figure 27. Western blot of α-UNC119 antisera against mouse embryo extracts. A) A ~70K product is detected by a 1:100 dilution of OF3 antiserum. Specific signal attenuation is observed when OF3 is pre-absorbed with high concentrations of exogenous UNC119-GST protein (7.1×10^{-5} μg protein per μl of antiserum used). A smaller product that migrates at a rate consistent with the known UNC119 size of ~33K is also very faintly detected, however it does not appear to be specifically out-competed by pre-absorption with UNC119-GST. OF5 antiserum faintly recognizes several products non-specifically. The membrane shown was exposed for 1 minute. B) A saturating exposure (2.25 hours) of the same membrane confirms a reduction of the ~70K signal but not of the ~33K signal when OF3 is pre-absorbed with UNC119. OF5 detection of several large bands is confirmed. The expected ~33K signal for native UNC119 is not observed using this antiserum. 5μl per lane of total mouse E14.5 protein extract was size fractionated and transferred to the membrane shown. PI= pre-immune serum.

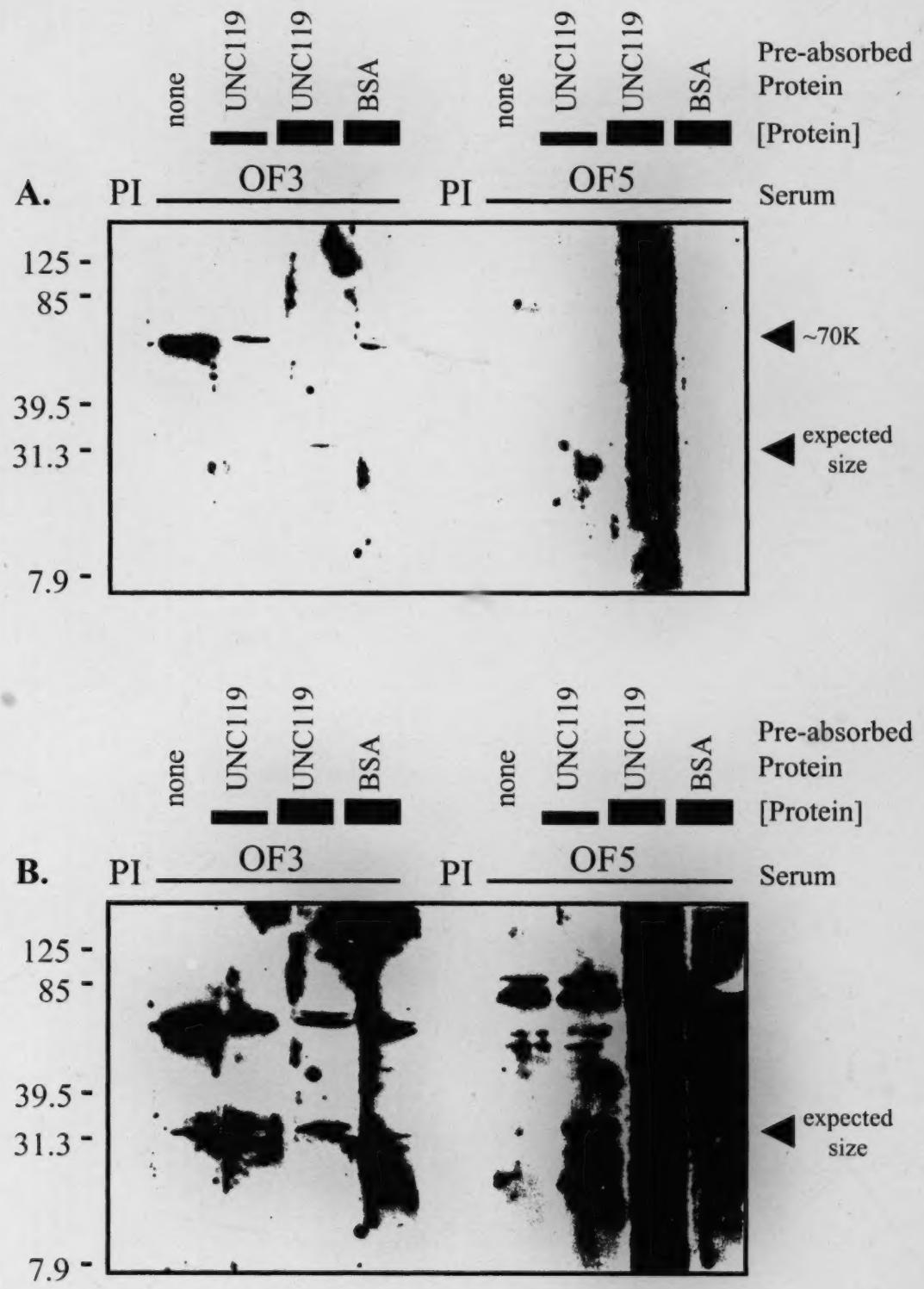
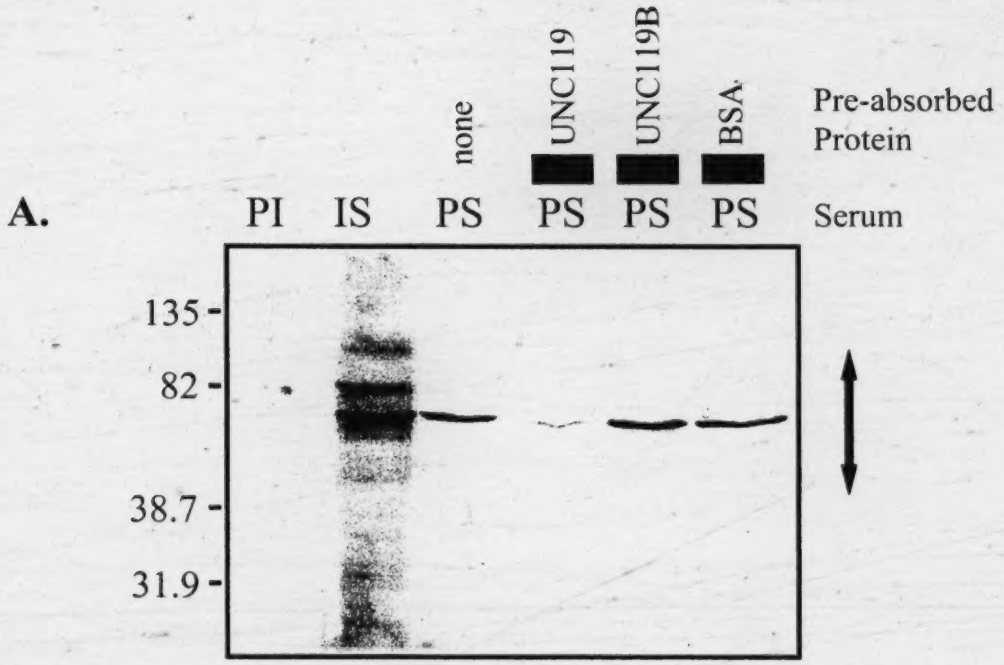

A.
PI
OF3
PI
OF5
Serum
none
UNC119
UNC119
BSA
none
UNC119
UNC119
BSA
Pre-absorbed Protein
[Protein]
125
85
39.5
31.3
7.9
~70K
expected size
B.
PI
OF3
PI
OF5
Serum
none
UNC119
UNC119
BSA
none
UNC119
UNC119
BSA
Pre-absorbed Protein
[Protein]
125
85
39.5
31.3
7.9
expected size

Figure 28. Western blot and densitometry analysis of α-UNC119 purified from OF3. A) A ~70K product is detected by a 1:100 dilution of purified OF3 antiserum after a short (5 second) exposure. Specific signal reduction is observed when purified OF3 is pre-absorbed with high concentrations of exogenous UNC119-GST (0.11μg protein per μl of antiserum). 5μl of E13.5-E14.5 total mouse embryo protein extract was size fractionated per lane. B) The proportion of signal attenuation by pre-absorption to exogenous UNC119 has been determined by densitometrical analysis using ImageJ software (Rasband 2002). A densitometry plot of the relative differences in signal intensity between each lane of the blot shown in A. Units of absorption have been assigned arbitrarily, and position refers to the region of the blot analyzed. The double-headed arrow indicates the region analysed, and the position to the left corresponds to the top of the analyzed region. IS= immune serum; PI= pre-immune serum; PS= purified immune serum. For the plot, UNC119 corresponds to PS pre-absorbed with UNC119-GST; UNC119B corresponds to PS pre-absorbed with UNC119B-GST; BSA corresponds to PS pre-absorbed with BSA.


A.
PI
IS
PS
PS
PS
PS
Serum
none
UNC119
UNC119B
BSA
Pre-absorbed Protein
135
82
38.7
31.9

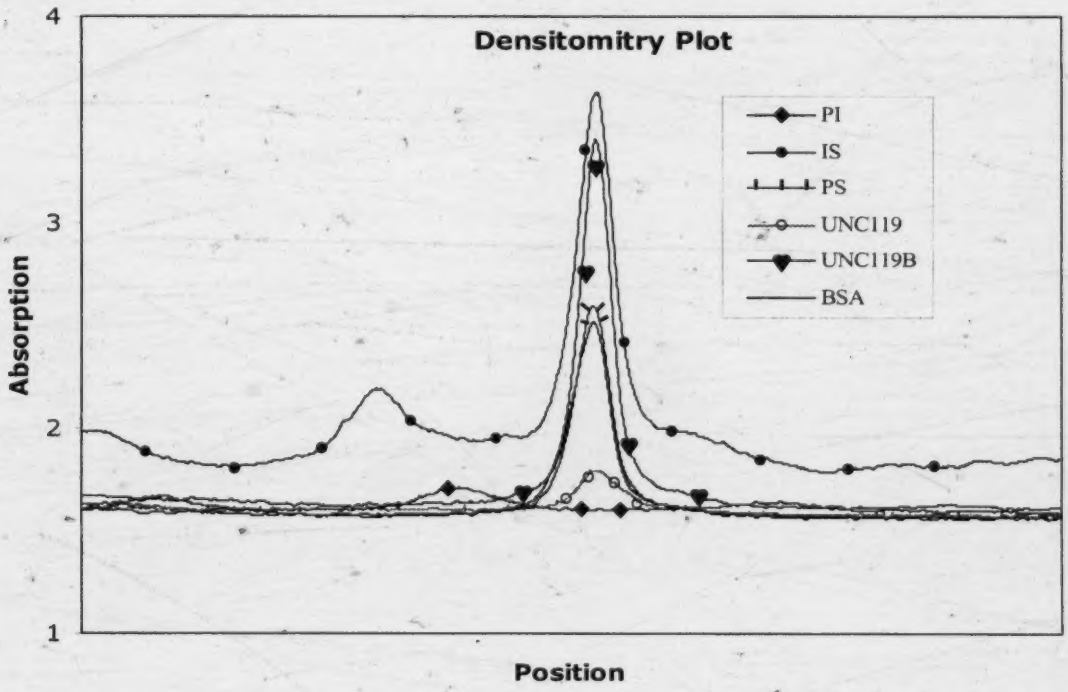
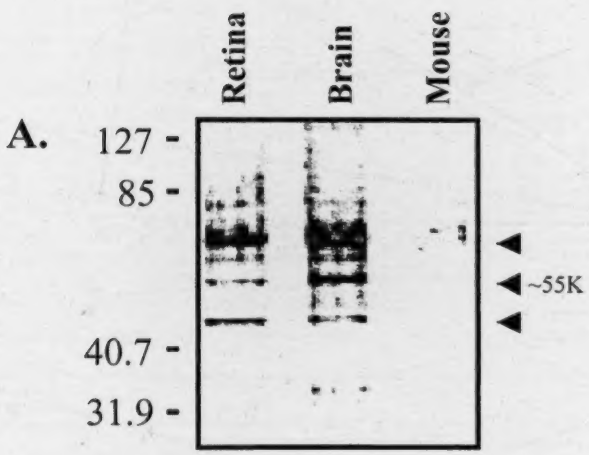

B.
Densitomitry Plot
PI
IS
PS
UNC119
UNC119B
BSA
Absorption
Position
4
3
2
1

Figure 29. Characterization of products recognized by α–UNC119 purified from OF3. A) Several products are eluted from α–UNC119 (purified from OF3) conjugated beads following incubation with adult bovine brain or retinal protein extracts, or with total mouse E14.5 extract. 5μl of each immuno-precipitate was fractionated on the gel shown. B) In each extract, 3 of the immuno-precipitated products are detected by a 1:100 dilution of purified α–UNC119 OF3 antibody. The ~55K product migrates at the rate expected for the IgG heavy chain of the antibody. The larger and smaller ones migrate at rates consistent with the specifically out-competed bands in the mouse and retina extracts respectively.


A.
Retina
Brain
Mouse
127
85
40.7
31.9
~55K

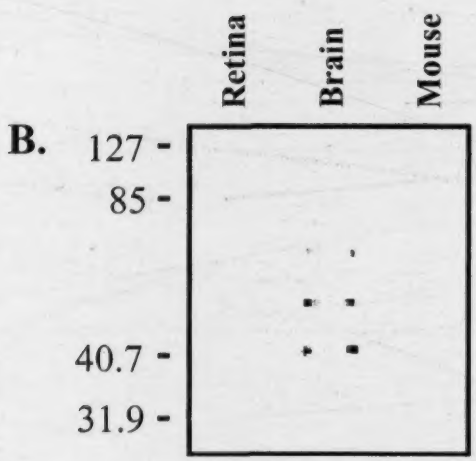

B.
Retina
Brain
Mouse
127
85
40.7
31.9

Figure 30. Western blot of α-UNC119B against mouse embryo protein. A) Products that migrate at ~80K and ~120K are faintly detected by 1:500 dilutions of 11-5 immune serum and purified 11-5 antiserum respectively. The blot shown was exposed for 15 minutes. B) More concentrated (1:100 dilution) applications of antiserum detect 5 products on blots exposed for several hours. Specific signal attenuation of 2 products (arrowheads) is observed when purified 11-5 is pre-absorbed with high concentrations of UNC119B-GST (7.0×10^{-2} μg protein per μl of antiserum used). 2μl aliquots of E14.5 total mouse protein extract were size fractionated and transferred to the membranes shown. IS= immune serum; PI= pre-immune serum; PS= purified immune serum.

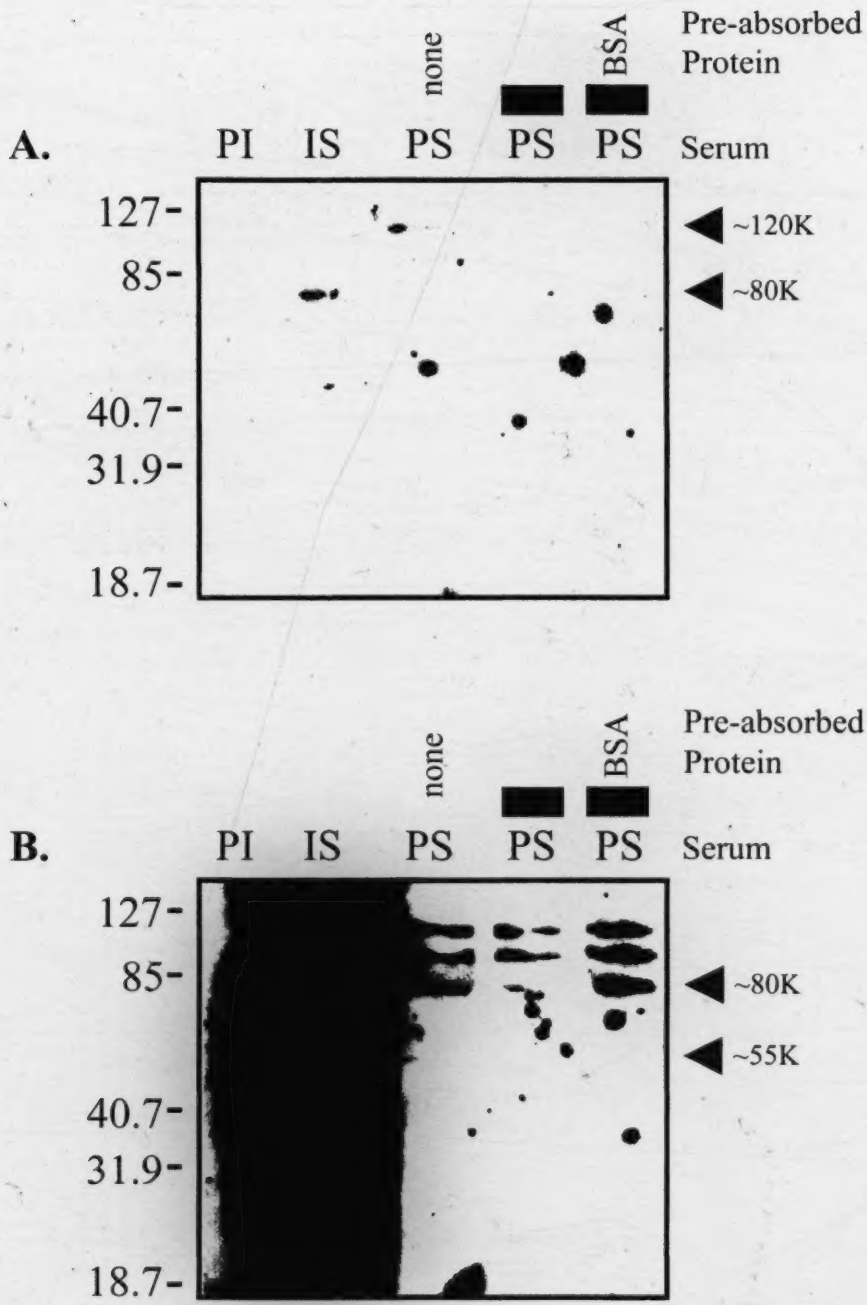

A.
Pre-absorbed Protein
none
BSA
PI
IS
PS
PS
PS
Serum
127-
85-
40.7-
31.9-
18.7-
~120K
~80K
B.
Pre-absorbed Protein
none
BSA
PI
IS
PS
PS
PS
Serum
127-
85-
40.7-
31.9-
18.7-
~80K
~55K

Figure 31. Western blot of α–UNC119B immuno-precipitates. The products immuno-precipitated from E14.5 total mouse extract, and from adult bovine brain and retinal protein extracts are recognized by a 1:100 dilution of α–UNC119B purified from 11-5 antiserum. 5μl of each immuno-precipitate was fractionated on the gel shown. Mass spectrometry indicates that the ~55K product corresponds to the IgG heavy chain of the antibody. The peptide summary report of the mass spectrometry analysis is demonstrated in Appendix 5.

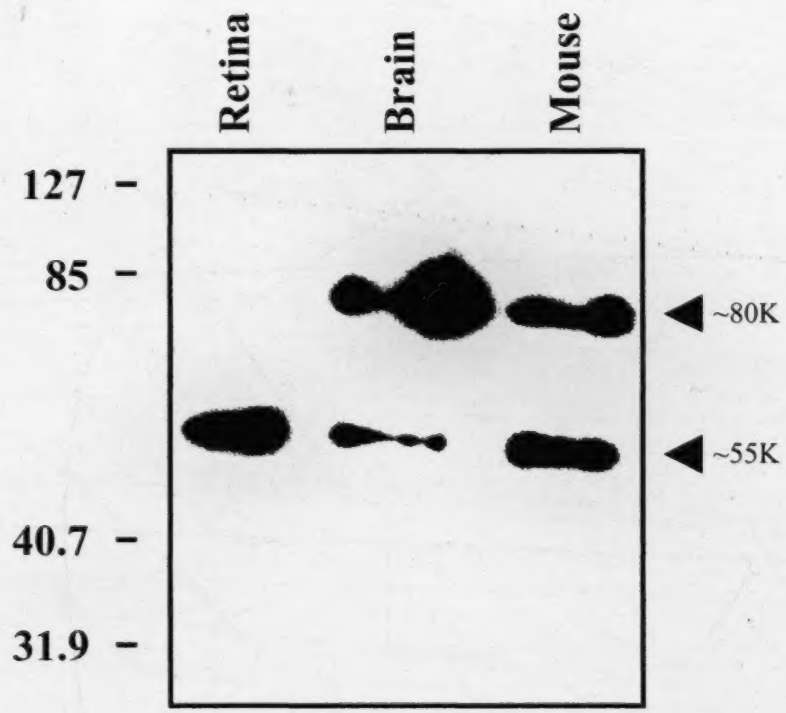

Retina
Brain
Mouse
127
85
40.7
31.9
~80K
~55K

Figure 32. Un-purified α-UNC119 immuno-localization on mouse cryo-sections. Un-purified α-UNC119 OF3 and OF5 staining was performed on frozen sagittal sections of E14.5 mouse embryos. A 1:500 dilution of OF3 antiserum (B) detects rows of cells (arrows) dorsal and anterior to the developing vertebrae compared to the same dilution of pre-immune OF3 serum applied to a nearly serial section (A). A 1:250 dilution of OF5 antiserum (D) detects small clusters of cells (arrows) immediately dorsal of each developing vertebra compared to a parallel application of its respective pre-immune serum (C). The clusters of cells may correspond to spinal nerves (Bagnall, personal communication). For each section, anterior is to the bottom and dorsal is to the right. The data required to determine the total magnification of the images shown has been lost.

Pre-immune serum **Immune serum**

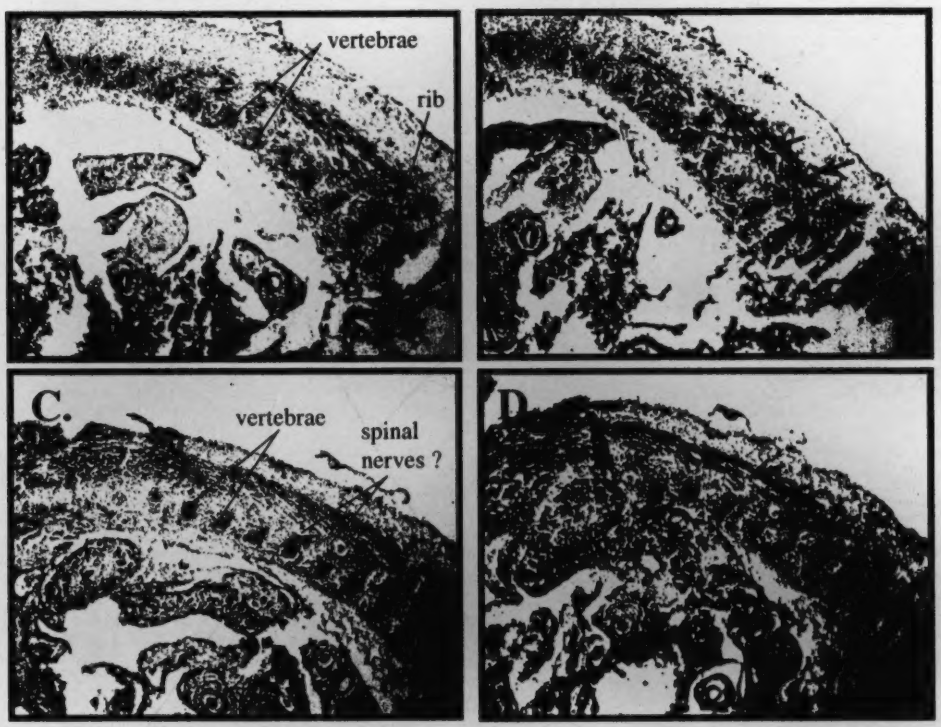


Figure 33. Immuno-localization of UNC119 on mouse embryo sections. Purified OF3 antiserum was applied at a 1:50 dilution to randomly oriented sections of E8.5 (A, B) and of E10.5 (C, D) mouse embryos *in utero*, and to E14.5 (E, F) para-sagittal sections. Purified antiserum was pre-absorbed overnight with UNC119-GST (0.1μg protein per μl of antiserum) for B, D, and F. No detection of UNC119 is apparent at E8.5. At E10.5 and E14.5, low levels of signal widely distributed throughout the sections appear more prominent on non pre-absorbed treated sections. Approximate magnifications are as follows: 55X for A and B; 20X for C and D; 7X for E and F.

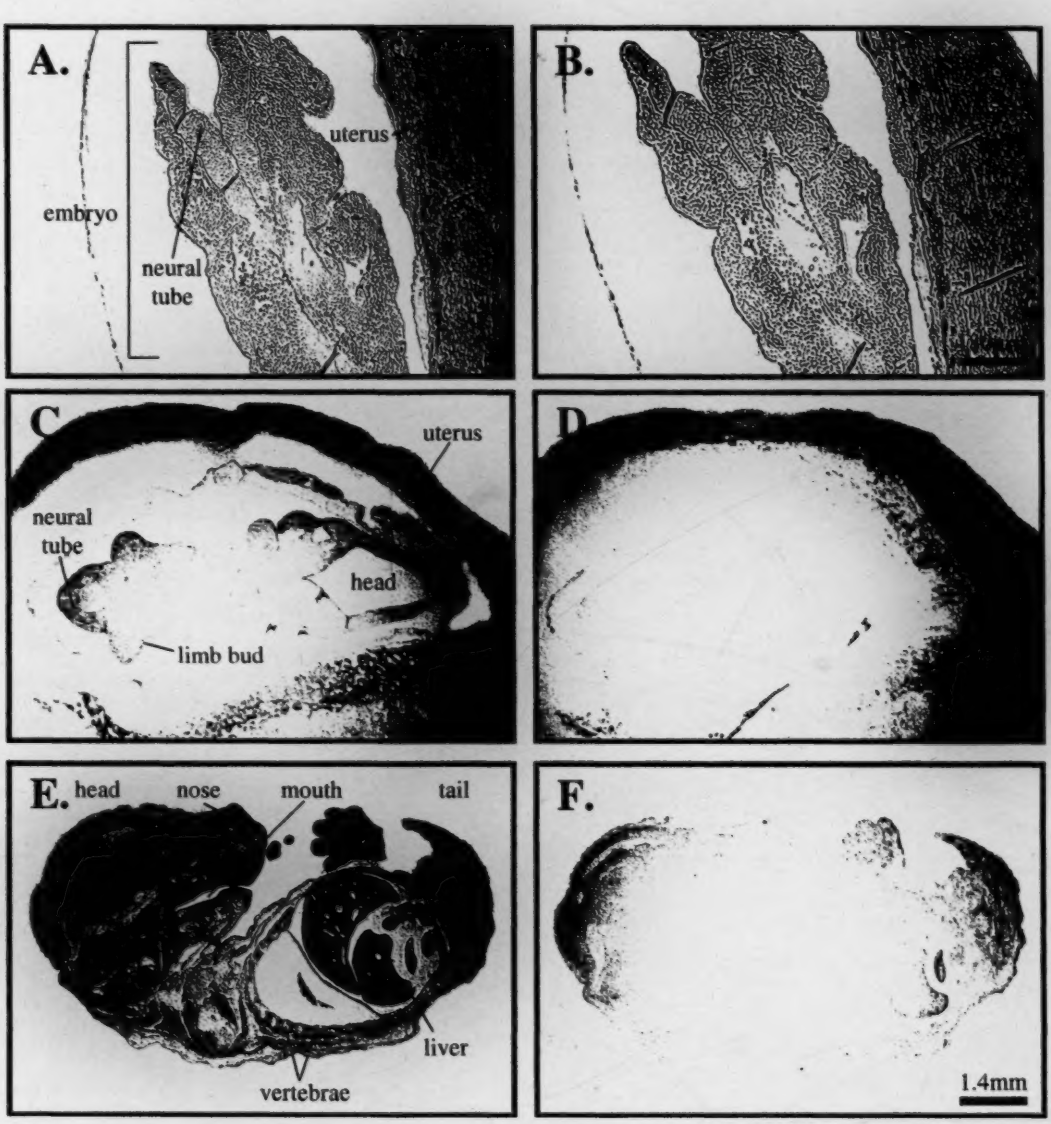
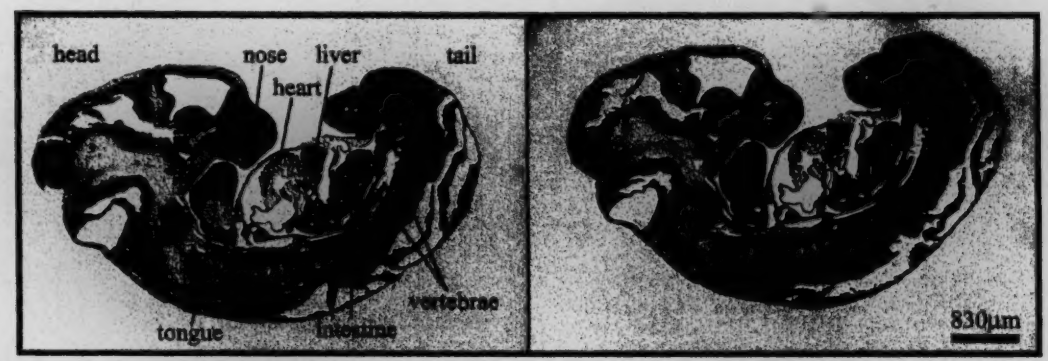
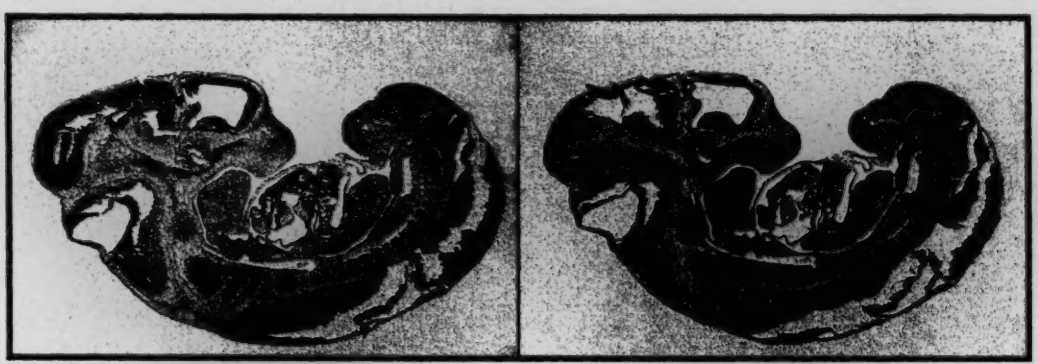

A.
uterus
embryo
neural tube
B.
C
uterus
neural tube
head
limb bud
D
E.
head
nose
mouth
tail
liver
vertebrae
F.
1.4mm

Figure 34. Immuno-localization of UNC119B on mouse embryo sections. A) A 1:250 dilution of purified 11-5 antiserum generates a faint widely distributed signal with some enrichment in the neural tube. B) A 1:250 dilution of un-purified pre-immune 11-5 antiserum generates a less pronounced staining pattern that is similar to the one observed in A. C) A 1:25 dilution of purified pre-immune 11-5 serum (which has an equivalent epitope titre to the 1:250 dilutions of pre-immune 11-5 and of purified 11-5) yields no discernible signal. All treatments were applied to duplicate serial sagittal sections of E12.5-E13.5 mouse embryos. Approximate magnification is 12X.

A.


head
nose
liver
tail
heart
vertebrae
tongue
intestine
830µm

B.



C.

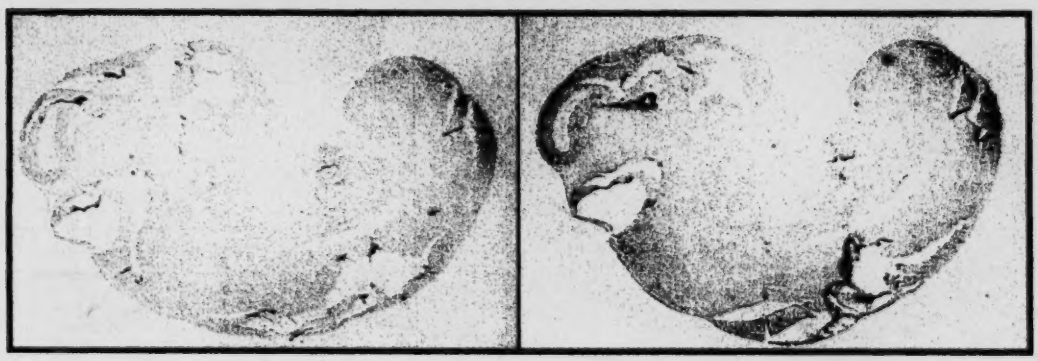


Figure 35. PC-12 cells produce UNC119 and UNC119B. 1:100 dilutions of α-UNC119 (purified from OF3) and of α-UNC119 (purified from 11-5) each strongly detect a single band in protein extracts from an un-induced PC-12 rat pheochromocytoma cell line. Longer exposures of the same membrane (shown to the right) confirm that there is no cross-reaction between the 2 proteins. 3µl aliquots of protein extract were size fractionated on the membrane shown. h= hour; sec= second.

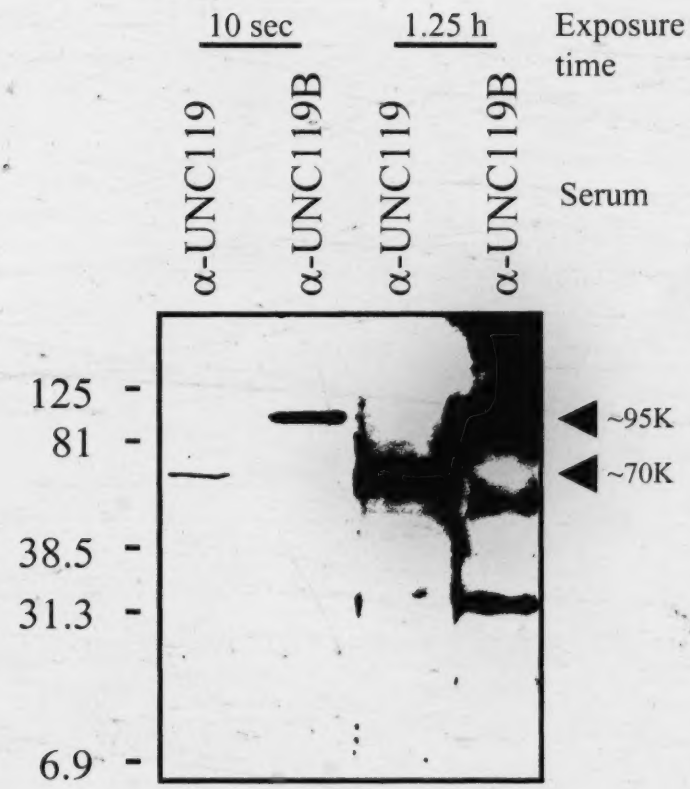

10 sec
1.25 h
Exposure time
α-UNC119
α-UNC119B
α-UNC119
α-UNC119B
Serum
125
81
38.5
31.3
6.9
~95K
~70K

Figure 36. Protein gel of UNC119-GST binding proteins. The left lane of the gel demonstrated corresponds to a 0.1μg aliquot of UNC119-GST fusion protein. The right lane corresponds to a 30μl aliquot of the eluted fraction from UNC119-GST coated beads following incubation with a non-denatured E14.5 total mouse embryo protein lysate. All bands present only in the right lane correspond to potential UNC119-binding proteins.

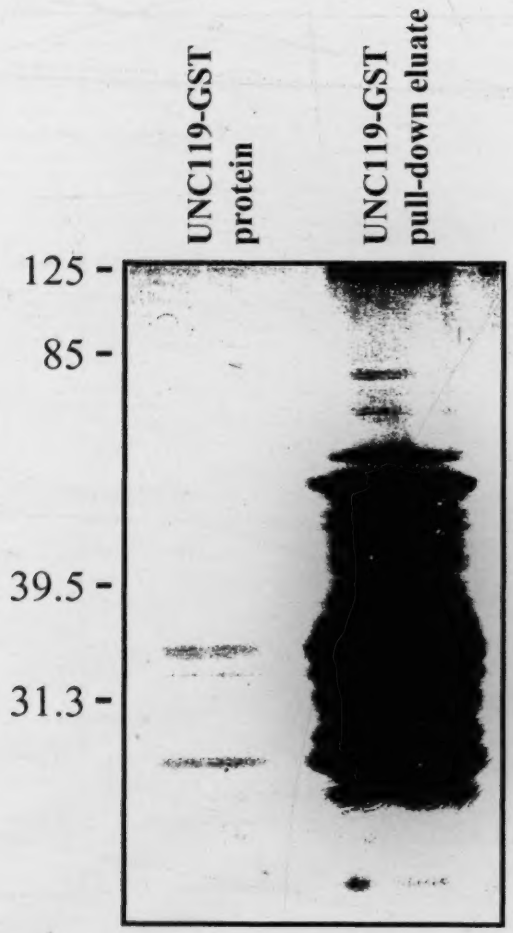

UNC119-GST protein
UNC119-GST pull-down eluate
125
85
39.5
31.3

Figure 37. Protein gel of mouse embryo proteins that immuno-precipitate with α-UNC119. The A/Ab lane (left) represents eluted proteins from un-purified α-UNC-119 (from rabbit OF3) coated beads that were not incubated with protein lysate. The IP lane represents a 4μl aliquot of an eluate from α-UNC119 OF3 coated protein A beads following incubation with non-denatured E14.5 total mouse protein. The 2 lanes to the right represent BSA loading standards of 0.2μg and 0.1μg respectively. The arrowheads denote bands that were extracted from the gel and analyzed by mass spectrometry. The most relevant portions of the protein summary reports for the extracted bands are located in Appendices 6, 7, and 8 respectively.

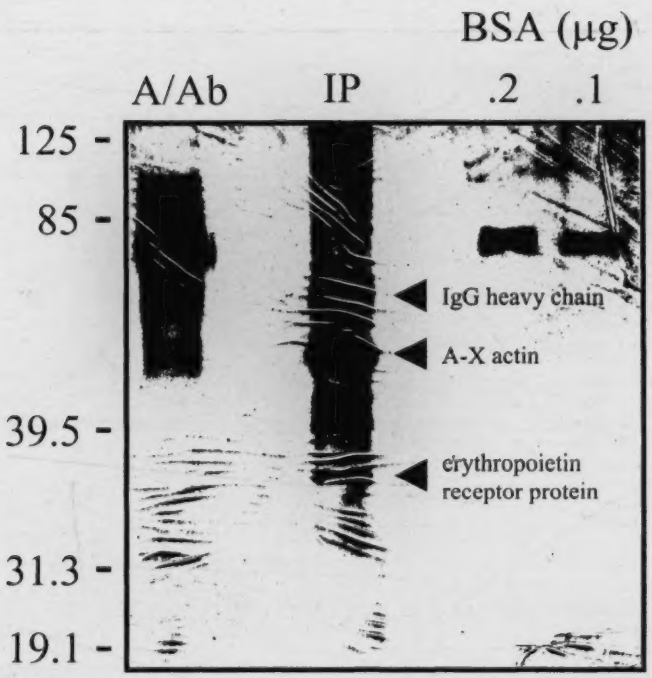

BSA (µg)
A/Ab
IP
.2
.1
125
85
39.5
31.3
19.1
IgG heavy chain
A-X actin
erythropoietin
receptor protein

Figure 38. The UNC-119 protein family. An amino acid alignment of the members of the UNC-119 protein family highlighting important features. The circle highlights a potential tyrosine nitration modification site. The thick line represents the conserved SH2 domain. The thin line represents the human UNC119 SH3 domain. Arrows highlight tyrosine residues that are potential substrates for Src tyrorine kinase phosphorylations. The sequences were aligned using CLUSTAL W (1.74) with the default settings (Thompson *et al.* 1994). Residues identical or similar among more than half of the sequences are highlighted in black and gray, respectively.

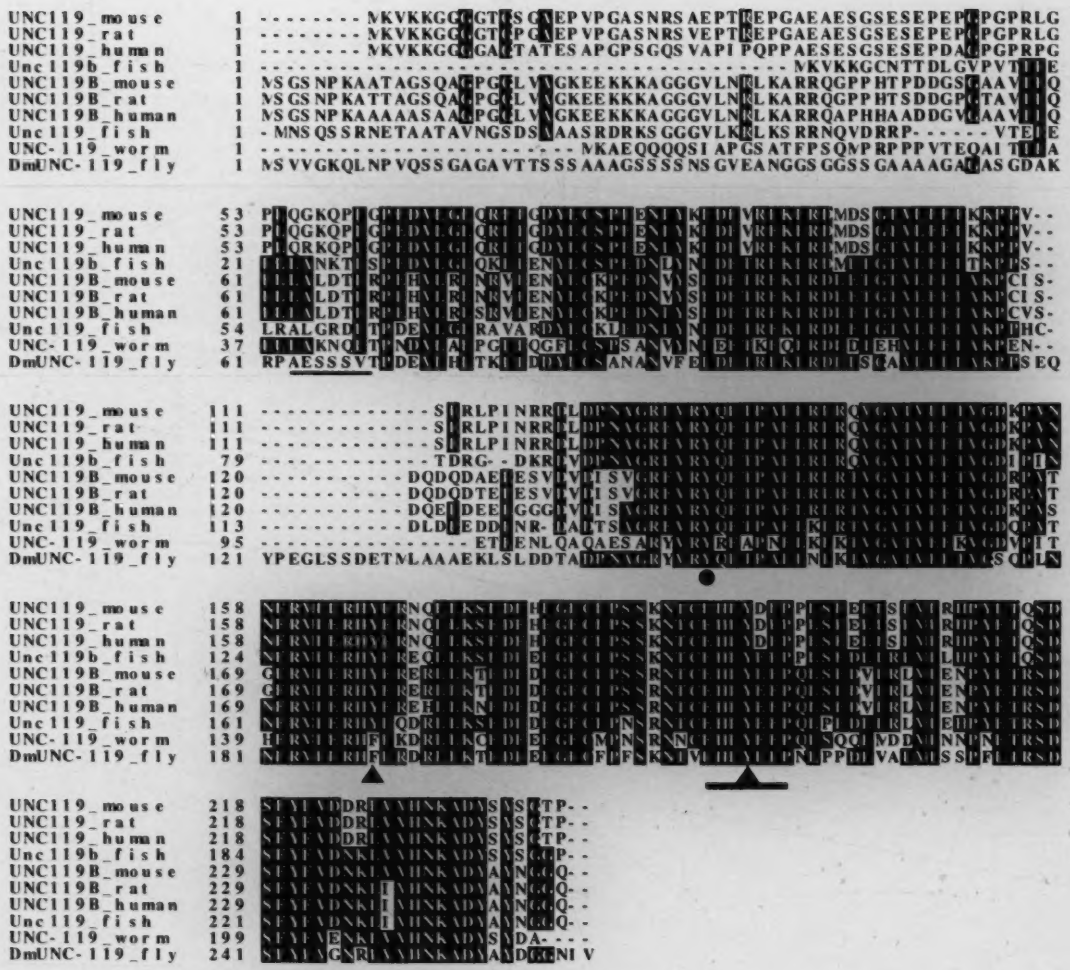
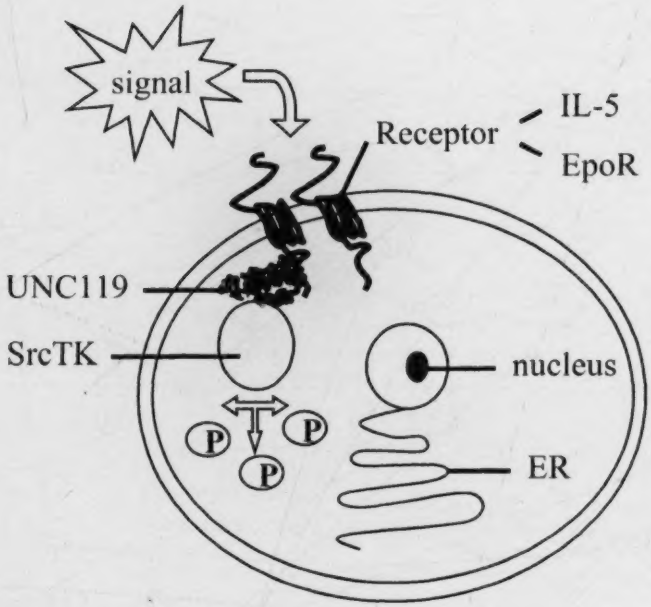
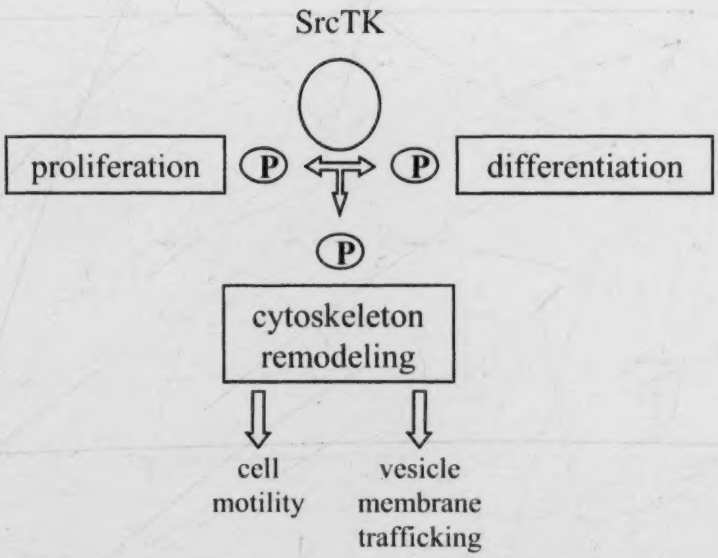

UNC119_mouse
UNC119_rat
UNC119_human
Unc119b_fish
UNC119B_mouse
UNC119B_rat
UNC119B_human
Unc119_fish
UNC-119_worm
DmUNC-119_fly

Figure 39. A wild speculation of UNC-119 function. A) UNC-119 acts as an adapter molecule in Src-mediated signal transduction cascades. It transmits extra-cellular signals from receptors of the cytokine superfamily to cytoplasmic SrcTKs. UNC-119 is important for the specificity of the cellular changes induced in response to various signals because of its bipartite structure. The highly conserved C-terminal is responsible for the signal transmission, while the differing protein binding capacities of the poorly conserved N-terminals are important for determining the specific kinases activated by a particular signal. B) Once activated, SrcTKs phosphorylate several types of cytosolic proteins, inducing cascades of changes that eventually result in a variety of cellular responses. These responses include proliferation, differentiation, and cytoskeletal remodeling, a process critical for cell plasticity, motility, and membrane organization. Defects in src-mediated signaling could therefore explain how UNC-119 is important for eosinophil proliferation, vesicle trafficking in photoreceptor cells, and axonal outgrowth and synapse establishment in invertebrates. EpoR= erythropoietin receptor; ER= endoplasmic reticulum; IL-5= interlukin 5 receptor; SrcTK= src tyrosine kinase.

A.


signal
Receptor
IL-5
EpoR
UNC119
SrcTK
nucleus
P
P
P
ER

B.


SrcTK
proliferation
P
P
differentiation
P
cytoskeleton remodeling
cell motility
vesicle membrane trafficking

Bibliography

Ahern-Djamali S, Bachmann C, Hua P, Reddy S, Kastenmeier A, Walter U and F Hoffmann (1999) Identification of Profilin and Src homology 3 domains as binding partners for *Drosophila* Enabled. *Proc Natl Acad Sci U S A* 96:4977-4982.

Alberts B, Bray D, Lewis J, Raff M, Roberts K and J Watson (1994) *Molecular Biology of the Cell*, 3rd ed., Garland Publishing, Inc., New York NY.

Alexandropoulos K, Cheng G and D Baltimore (1995) Proline-rich sequences that bind to Src homology 3 domains with individual specificities. *Proc. Natl. Acad. Sci. U S A* 92:3110-3114.

Ausubel F, Brent R, Kingston R, Moore D, Smith J, Seidman J and K Struhl (1987) *Current Protocols in Molecular Biology*. Greene Publishing Associates, New York NY.

Ayscough K (1998) In vivo functions of actin-binding proteins. *Curr Opin Cell Biol* 10:102-111.

Bartolini F, Bhamidipati A, Thomas S, Schwahn U, Lewis S and N Cowan (2002) Functional overlap between Retinitis Pigmentosa 2 protein and the tubulin-specific chaperone cofactor C. *J Biol Chem* 277:14629-14634.

Blom N, Gammeltoft S and S Brunak (1999) Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* 294:1351-1362.

Carlsson L, Nystrom L, Sundkvist I, Markey F and U Lindberg (1977) Actin polymerizability is influenced by Profilin, a low molecular weight protein in non-muscle cells. *J Mol Biol* 115:465-83.

Cen O, Gorska M, Stafford S, Sur S and R Alam (2003) Identification of UNC119 as a novel activator of Src-type tyrosine kinases. *J Biol Chem* 278:8837-8845.

Chen S, Wang Q, Nie Z, Sun H, Lennon G, Copeland N, Gilbert D, Jenkins N and D Zack (1997) Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron* 19:1017-1030.

Clark J, Moore L, Krasinskas A, Way J, Battey J, Tamkun J and A Kahn (1993) Selective amplification of additional members of the ADP-Ribosylation Factor (ARF) family: Cloning of additional human and *Drosphila* ARF-like genes. *Proc Natl Acad Sci U S A* 90:8952-8956.

Combet C, Blanchet C, Geourjon C and G Deleage (2000) NPS@: network protein sequence analysis. *Trends Biochem Sci* 25:147-150.

Cooke J, Nowak M, Boerlijst M and J Maynard-Smith (1997) Evolutionary origins and maintenance of redundant gene expression during metazoan development. *Trends Gene* 13:360-364.

Corpet F, Gouzy J and D Kahn (1998) The ProDom database of protein domain families. *Nucleic Acids Res* 26:323-326.

Creighton T (1988) Disulphide bonds and protein stability. *Bioessays* 8:57-63.

D'Andrea R and T Gonda (2000) A model for assembly and activation of the GM-CSF, IL-3 and IL-5 receptors: insights from activated mutants of the common β subunit. *Exp Hematol* 28:231-243.

Davidson S, Jonas A, Clayton D and J George (1998) Stabilization of α-synuclein secondary structure upon binding to synthetic membranes. *J Biol Chem* 273:9443-9449.

Davis C, Homyard D, Millen K and A Joyner (1991) Examining pattern formation in mouse, chicken, and frog embryos with an *En*-specific antiserum. *Development* 2:287-298.

Deutsch U, Dressler G and P Gruss (1988) *Pax1*, a member of a paired box homologous mouse gene family, is expressed in segmented structures during development. *Cell* 53:617-625.

Donaldson J and C Jackson (2000) Regulators and effectors of the ARF GTPases. *Curr Opin Cell Biol* 12:475-482.

Emanuelsson O, Nielsen H, Brunak S and G von Heijne (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300:1005-1016.

Evans C, Pierce A, Winter S, Spooncer E, Heyworth C and A Whetton (1999) Activation of Granulocyte-Macrophage Colony-Stimulating Factor and Interleukin-3 Receptor subunits in a multipotential hematopoietic progenitor cell line leads to differential effects on development. *Blood* 94:1504-1514.

Fariselli P, Riccobelli P and R Casadio (1999) Role of evolutionary information in predicting the disulfide-bonding state of cysteine in proteins. *Proteins* 36:340-346.

Fisher J (2003) Erythropoietin: Physiology and pharmacology update. *Exp Biol Med* 228:1-14.

Florio S, Prusti R and J Beavo (1996) Solubilization of membrane-bound rod phosphodiesterase by rod phosphodiesterase recombinant δ subunit. *J Biol Chem* 271:24036-24047.

Force A, Lynch M, Pickett B, Amores A, Yan Y, and J Postlethwait (1998) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531-1545.

Furukawa T, Morrow E and C Cepko (1997) *Crx*, a novel *Otx*-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell* 91:531-541.

Gardner C and K Barald (1992) Expression patterns of engrailed-like proteins in the chick embryo. *Dev Dyn* 193:370-388.

Geijsen N, Koenderman L and P Coffer (2001) Specificity in cytokine signal transduction: Lessons learned from the IL-3/IL-5/GM-CSF receptor family. *Cytokine Growth Factor Rev* 12:19-25.

Gertler F, Comer R, Juang J, Ahern S, Clark M, Liebl E and F Hoffmann (1995) Enabled, a dosage-sensitive suppressor of mutations in the Drosophila Abl tyrosine kinase, encodes an Abl substrate with SH3 domain-binding properties. *Genes Dev* 9:521-533.

Gilbert S (2000) *Developmental Biology*, 6th ed., Sinauer Associates, Inc., Sunderland MA.

Gillespie P, Prusti R, Apel E and J Beavo (1989) A soluble form of bovine rod photoreceptor phosphodiesterase has a novel 15-kDa subunit. *J Biol Chem* 264:12187-12193.

Greenwald I (1985) Lin-12, a nematode homeotic gene, is homologous to a set of mammalian proteins that includes epidermal growth factor. *Cell* 43:583-590.

Gupta R, Hansen J and S Brunak (2003) *Identifying intracellular O-(beta)-GlcNAc 'yin-yang' switches in the available human proteome.* Manuscript in preparation.

Halder G, Callaerts P and W Gehring (1995) Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. *Science* 267:1788-1792.

Hansen J, Lund O, Tolstrup N, Gooley A, Williams K and S Brunak (1998) NetOglyc: Prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconj J* 15:115-130.

Hansen J, Lund O, Rapacki K and S Brunak (1997) O-glycbase version 2.0 - A revised database of O-glycosylated proteins. *Nucleic Acids Res* 25:278-282.

Hansen J, Lund O, Engelbrecht J, Bohr H, Nielsen J, Hansen J and S Brunak (1995) Prediction of O-glycosylation of mammalian proteins: Specificity patterns of UDP-GalNAc-polypeptide N-acetylgalactosaminyltransferase. *Biochem J* 308:801-813.

Hanzal-Bayer M, Renault L, Roversi P, Wittinghofer A and R Hillig (2001) The complex of Arl2-GTP and PDEδ: From structure to function. *EMBO J* 21:2095-2106.

Hawley A (2002) Characterization of Dmunc-119, the *Drosophila* homologue of *C. elegans unc-119*. M.Sc. Thesis. McMaster University, Hamilton ON.

Higashide T and G Inana (1999) Characterization of the gene for HRG4 (UNC119), a novel photoreceptor synaptic protein homologous to Unc-119. *Genomics* 57:446-450.

Higashide T, McLaren M. and G Inana (1998) Localization of HRG4, a photoreceptor protein homologous to Unc-119, in ribbon synapse. *Invest Opthalmol Vis Sci* 39:690-698.

Higashide T, Murakami A, McLaren M and G Inana (1996) Cloning of the cDNA for a novel photoreceptor protein. *J Biol Chem* 271:1797-1804.

Hill M and P Gunning (1993) Beta and gamma actin mRNAs are differentially located within myoblasts. *J Cell Biol* 122:825-832.

Hofmann K, Bucher P, Falquet L and A Bairoch (1999) The PROSITE database, its status in 1999. *Nucleic Acids Res* 27:215-219.

Hogan B *et al.* (1994) *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York NY.

Huet C (1998) Production of Polyclonal Antibodies in Rabbits. In *Cell Biology: A Laboratory Handbook*, 2nd ed., vol. 2. Academic Press Ltd., New York NY.

Ischiropoulos H (1998) Biological tyrosine nitration: a pathophysiological function of nitric oxide and reactive oxygen species. *Arch Biochem Biophys* 356:1-11.

Jean D, Ewan K and P Gruss (1998) Molecular regulators involved in vertebrate eye development. *Mech Dev* 76:3-18.

Joyner A and G Martin (1987) *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila engrailed* gene: Expression during embryogenesis. *Genes Dev* 1:29-38.

Kaufman M (1995) *The Altas of Mouse Development*. Academic Press Ltd., San Diego CA.

Kaufman M and J Bard (1999) *The Anatomical Basis of Mouse Development*. Academic Press Ltd., San Diego CA.

Ketteler R, Heinrich A, Offe J, Becker V, Cohen J, Neumann D and U Klingmüller (2002) A functional green fluorescent protein-tagged erythropoietin receptor despite physical separation of JAK2 binding site and tyrosine residues. *J Biol Chem* 277:26547-26552.

Knobel K, Davis W, Jorgensen E and M Bastiani (2001) UNC-119 suppresses axon branching in *C. elegans*. *Development* 128:4079-4092.

Kobayashi A, Kubota S, Mori N, McLaren M and G Inana (2003) Photoreceptor synaptic protein HRG4 (UNC119) interacts with ARL2 via a putative conserved domain. *FEBS Lett* 534:26-32.

Kobayashi A, Higashide T, Hamasaki D, Sakuma H, Kubota S, An W, Fujimaki T, McLaren M, Weleber R and G Inana (2000) HRG4 (UNC119) mutation found in cone-rod dystrophy causes retinal degeneration in a transgenic model. *Invest Opthalmol Vis Sci* 41:3268-3277.

Kozak M (2000) Do the 5' untranslated domains of human cDNAs challenge the rules for initiation of translation (or is it vice versa)? *Genomics* 70:396-406.

Kozak M (1996) Interpreting cDNA sequences: some insights from studies on translation. *Mamm Genome* 7:563-574.

Kubota S, Kobayashi A, Mori N, Higashide T, McLaren M and G Inana (2002) Changes in retinal synaptic proteins in the transgenic model expressing a mutant. *Invest Opthalmol Vis Sci* 43:308-313.

Lee S and D Jacobs (1999) Expression of *Distal-less* in molluscan eggs, embryos, and larvae. *Evol Dev* 1:172-179.

Li N, Florio S, Pettenati M, Rao N, Beavo J and W Baehr (1998) Characterization of human and mouse rod cGMP phosphodiesterase δ subunit (PDE6D) and chromosomal localization of the human gene. *Genomics* 49:76-82.

Lin C and D Sargan (2001) Generation and analysis of canine retinal ESTs: Isolation and expression of retina-specific gene transcripts. *Biochem Biophys Res Commun* 282:394-403.

Linari M, Ueffing M, Manson F, Wright A, Meitinger T and J Becker (1999) The retinitis pigmentosa GTPase regulator, RPGR, interacts with the delta subunit of rod cyclic GMP phosphodiesterase. *Proc Natl Acad Sci U S A* 96:1315-1320.

Maduro M (1998) Characterization of the *unc-119* gene of *Caenorhabditis elegans*. Ph.D. Thesis. University of Alberta, Edmonton AB.

Maduro M, Gordon M, Jacobs R, and D Pilgrim (2000) The UNC-119 family of neural proteins is functionally conserved between humans, *Drosophila* and *C. elegans*. *J Neurogenet* 13:191-212.

Maduro M and D Pilgrim (1996) Conservation of function and expression of *unc-119* from two *Caenorhabditis* species despite divergence of non-coding DNA. *Gene* 183:77-85.

Maduro M and D Pilgrim (1995) Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* 141:977-988.

Manning A (1999) Characterization of *unc-119* homologues in the zebrafish, *Danio rerio*. M.Sc. Thesis. University of Alberta, Edmonton AB.

Marzesco A, Galli T, Louvard D and A Zahrahoui (1998) The rod cGMP phosphodiesterase δ subunit dissociates the small GTPase Rab13 from membranes. *J Biol Chem* 273:22340-22345.

Mitchell L, Mutchmor J and W Dolphin (1988) *Zoology*. The Benjamin/Cummings Publishing Company Inc., Menlo Park CA.

Mitchell P and R Tjian (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245:371-378.

Nasevicius A and S Ekker (2000) Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* 26:216-220.

Nie Z, Chen S, Kumar R and D Zack (1996) RER, an evolutionarily conserved sequence upstream of the rhodopsin gene, has enhancer activity. *J Biol Chem* 271:2667-2675.

Pomeroy S, Tamayo P, Gaasenbeek M, Sturla L, Angelo M, McLaughlin M, Kim J, Goumnerova L, Black P, Lau C, Allen J, Zagzag D, Olson J, Curran T, Wetmore C, Biegel J, Poggio T, Mukherjee S, Rifkin R, Califano A, Stolovitzky G, Louis D, Mesirov J, Lander E and T Golub (2002) Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* 415:436-442.

Postlethwait J, Yan Y, Gates M, Horne S and A Amores (1998) Vertebrate genome evolution and the zebrafish gene map. *Nat Genet* 18:345-349.

Radcliffe P, Vardy L and T Toda (2000) A conserved small GTP-binding protein Alp41 is essential for the cofactor-dependent biogenesis of microtubules in fission yeast. *FEBS Lett* 468:84-88.

Rane S and E Reddy (2002) JAKs, STATs and Src kinases in hematopoiesis. *Oncogene* 21:3334-3358.

Rasband W (2002) ImageJ software (version 1.29). (http://rsb.info.nih.gov/ij/). National Institute of Health, Bethesda MD.

Rechsteiner M and S Rogers (1996) PEST sequences and regulation by proteolysis. *Trends Biochem Sci* 21:267-271.

Relaix F and M Buckingham (1999) From insert eye to vertebrate muscle: Redeployment of a regulatory network. *Genes Dev* 13:3171-3178.

Rogers S, Wells R and M Rechsteiner (1986) Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis. *Science* 234:364-368.

Rost B, Fariselli P and R Casadio (1996) Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci* 5:1704-1718.

Rost B and C Sander (1993) Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* 232:584-599.

Rugh R (1968) *The Mouse. Its Reproduction and Development.* Burgess Publishing Company, Minneapolis MN.

Saitou N and M Nei (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425.

Sambrook J and D Russell (2001) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York NY.

Sander C and R Schneider (1991) Database of homology-derived structures and the structural meaning of sequence alignment. *Proteins* 9:56-68.

Schorpp M, Leicht M, Nold E, Hammerschmidt M, Haas-Assenbaum A, Weist W and T Boehm (2002) A zebrafish orthologue (*whnb*) of the mouse *nude* gene is espressed in epithelial compartment of the embryonic thymic rudiment. *Mech Dev* 118:179-185.

Sharer J, Shern J and H Valkenburgh (2002) ARL2 and BART enter mitochondria and bind the adenine nucleotide transporter. *Mol Biol Cell* 13:71-83.

Songyang Z, Shoelson S, Chaudhuri M, Gish G, Pawson T, Haser W, King F, Roberts T, Ratnofsky S, Lechleider R, *et al.* (1993) SH2 domains recognize specific phosphopeptide sequences. *Cell* 72:767-778.

Souza J, Giasson B, Chen Q, Lee V and H Ischiropoulos (2000) Dityrosine cross-linking promotes formation of stable α-synuclein polymers. *J Biol Chem* 275:18344-18349.

Souza J, Daikhin E, Yudkoff M, Raman C and H Ischiropoulos (1999) Factors determining the selectivity of protein tyrosine nitration. *Arch Biochem Biophys* 371:169-178.

Suter D and F Forscher (2000) Substrate-cytoskeletal coupling as a mechanism for the regulation of growth cone motility and guidance. *J Neurobiol* 44:97-113.

Swanson D, Chang J, Campochiaro P, Zack D and D Valle (1998) Mammalian orthologs of *C. elegans Unc-119* highly expressed in photoreceptors. *Invest Opthalmol Vis Sci* 39:2085-2094.

Thompson J, Gibson T, Plewniak F, Jeanmougin F and D Higgins (1997) The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876-4882.

Thompson J, Higgins D and T Gibson (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673-4680.

Tzafrir I, McElver J, Liu C, Yang L, Wu J, Martinez A, Patton D and D Meinke (2002) Diversity of TITAN functions in Arabidopsis seed development. *Plant Physiol* 128:38-51.

Van Valkenburgh H, Shern J, Sharer J, Zhu X and R Kahn (2001) ADP-ribosylation factors (ARFs) and ARF-like 1 (ARL1) have both specific and shared effectors. *J Biol Chem* 276:22826-22837.

Vaux D, Weissman I and S Kim (1992) Prevention of programmed cell death in *Caenorhabditis elegans* by human bcl-2. *Science* 258:1955-1957.

Wills Z, Marr L, Zinn K, Goodman C and D Van Vactor (1999) Profilin and the Abl tyrosine kinase are required for motor axon outgrowth in the *Drosophila* embryo. *Neuron* 22:291-299.

Zack D, Bennett J, Wang Y, Davenport C, Klaunberg B, Gearhart J and J Nathans (1991) Unusual topography of bovine rhodopsin promoter-lacZ fusion gene expression in transgenic mouse retinas. *Neuron* 6:187-99.

Appendix 1. Detailed protocol for whole mount *in situ* hybridization

DAY 1

- All solutions prepared are filter sterilized to remove dirt particles.
- **Rehydrate:** **70% EtOH/PBT (DEPC)** **1 X 5 min each, rocking at room temperature**
  50% EtOH/PBT (DEPC)
  25% EtOH/PBT (DEPC)
  1X DEPC PBT **2 X 5 min**
  - Use reagent grade 100% ethanol and filtered DEPC treated 1X PBS + 0.1% Tween (PBT).
  - Place embryos in sterile 50ml tubes using ~20X volume for each wash.
  - Decant solutions over a sterile dish and use DEPC-SDS rinsed forceps to handle embryos.
- **Bleach:** **6% H_2O_2 in DEPC PBT** **1h, rocking at room temperature**
  vent caps often to release pressure builup
  - Prepare fresh using concentrated (30%) bleach.
  - Incubate in a sterile 50ml tube no more than 50% full.
- **Rinse:** **1X DEPC PBT** **3 X 5 min each, rocking slow at room temperature**
  - Decant embryos into a sterile dish and remove bleach using sterile pipets. Wash directly in dish; embryos are extremely fragile and should not be handled.
- **Permeabilize:** **20µg/ml proteinase K in DEPC PBT, rocking at room temperature**
  Incubation time depends on embryo age (plug date = 0.5) and size:
  - **10.5 = 14 min** (~4mm)
  - **11.5 = 20 min** (~7mm)
  - **12.5 = 27 min** (~9mm)
  - Make proteinase K fresh in PBT from frozen 10mg/ml stock thawed on ice.
  - Permeabilize in a sterile petri dish containing ~20ml solution. Different age embryos should be placed in separate dishes. Remove solution promptly and proceed with washes using the same dish.
- **Rinse:** **2mg/ml glycine in DEPC PBT** **1 X 10 min, rocking at room temperature**
  1X DEPC PBT **2 X 5 min, rocking at room temperature**
  - Make first wash solution using glycine powder in DEPC PBT.
- **Fix:** **4% paraformaldehyde / 0.2% glutaraldehyde in DEPC PBT**
  1 X 20 min, rocking at room temperature
  - Make fresh from high (EM) grade fixatives (Electron Microscopy Science, Inc.).
  - EM paraformaldehyde can be diluted to 4% in DEPC PBT and frozen for later use.
- **Rinse:** **1X DEPC PBT** **4 X 5 min, rocking at room temperature**
- **Pre-hybridization:** **Hybridization solution** **2 X 1h (minimum), rotating at 68°C**
  Hybridization solution: **(50% formamide, 5X SSC pH4.5, 1% SDS, 50µg/ml yeast RNA, 50µg/ml heparin)**
  - Hybridization solution prepared using RNAse free reagents, stored at –20°C, and pre-warmed to 68°C before use.
  - Transfer embryos to sterile 2ml tubes using 1.2ml hybridization solution. Place 2ml tubes (containing embryos) in pairs in 50ml tubes to fit on the hybridization oven rotor.
- **Hybridization:** **0.5-1µg/ml DIG riboprobe in hybridization solution**
  hybridize overnight at 68°C
  - Add hybridization solution to probes and heat to 80°C for 3-5 min before incubation. Immediately place in hybridization oven to prevent cooling, and add to each tube 1 at a time.

DAY 2

- Solutions used in subsequent steps are not DEPC treated.
- **Wash #1:** **wash (solution #1)** **2 X 30 min, rotating at 68°C**
  wash solution #1: (50% formamide, 5XSSC pH4.5, 1%SDS)
  - Solution is stored at 4°C and pre-warmed before use.
  - Tubes are almost filled for each wash and all solutions are decanted over a sterile dish holding back embryos with cap to minimize handling.
- **Equilibrate:** **1:1 (wash solution #1:RNase buffer)** **10 min, rotating at 68°C**

RNase buffer **3 X 5 min, rotating at room temperature**

RNAse buffer: (0.5M NaCl, 10mM Tris-Cl pH7.5, 0.1% Tween)

- **RNAse treatment:** **100µg/ml RNase A** **2 X 30 min, rotating at 37°C**
  - Use DNase free RNase A (10mg/ml) prepared as described elsewhere (Sambrook and Russell 2001) diluted in RNAse buffer immediately before use.
- **Start α-DIG antibody pre-absorption (described below) during second RNase treatment.**
- **Rinse:** **RNase buffer** **1 X 5 min, rotating at room temperature**
- **Wash #2:** **wash (solution #2)** **1 X 5 min, rotating at roon temperature**
  2 X 30 min, rotating at 68°C

  wash solution #2: (50% formamide, 2X SSC pH4.5)
  - Solution stored at 4°C and pre-warmed before use.
- **Equilibrate:** **1X TBT** **3 X 5 min, rotating at room temperature**
  - TBT is filter sterilized 1X TBS supplemented with 0.1% Tween 20.
- **Block:** **10% mouse serum** **2.5h (at least), rotating at room temperature**
  - Mouse serum (100%) is heat inactivated 30 min at 70°C, preserved with 0.02% sodium azide and stored at 4°C. Dilute to 10% in TBT before use.
- **Antibody (α-DIG) incubation:** **1:2000 pre-absorbed α-DIG antibody rotating overnight at 4°C**
  - Antibody pre-absorbed in 1% mouse serum, 2mM levamisole, and 1% blocking reagent (Boeringer Mannheim, Inc.) in TBT. Pre-absorption procedure described later.

DAY 3

- **Wash:** **2mM levamisole in 1 X TBT** **3 X 5 min, rotating at room temperature**
  2mM levamisole in 1 X TBT **5 X 1.5h, rotating at room temperature**
  2mM levamisole in 1 X TBT **overnight, rotating at 4°C**
  - Add levamisole to TBT on day of use and keep at 4°C between washes.

DAY 4

- Transfer embryos to 4ml glass tubes using curved forceps.
- **Equilibrate:** **AP developing solution** **3 X 5 min, rotating at room temperature**

  AP solution: (100mM Tris-Cl pH9.5, 100mM NaCl, 50mM $MgCl_2$, 2mM levamisole, 0.1% Tween 20)
  - Make fresh, keep on ice and use ~1.5ml per wash.
- **Develop:** **NBT/BCIP development mix**

  Development mix: 5µl NBT and 3.75 µl BCIP (Promega, Inc.) **per ml of AP solution**
  Incubate at room temperature in dark without shaking until colour optimal
  - Check colour development every 15-30 min.
- **Wash:** **20mM EDTA in PBT** **2 X 5min, rotating at room temperature**
  4-5 X 1h, rotating at room temperature
  overnight, rotating at room temperature
  - Use ~2ml per wash.

DAY 5

- **Post-fix:** **4% paraformaldehyde in PBT** **1h, rotating at room temperature**
- **Rinse:** **1 X PBT** **2 X 5 min, rotating at room temperature**
- **Clear :** **Dehydrate/rehydrate embryos using a methanol gradient:**
  - **25% MeOH in PBT** **1 X 5 min each, rotating at room temperature**
  - **50% MeOH in PBT**
  - **75% MeOH in PBT**
  - **100% MeOH in PBT** **2 X 5 min, rotating at room temperature**
  - **75% MeOH in PBT** **1 X 5 min each, rotating at room temperature**
  - **50% MeOH in PBT**
  - **25% MeOH in PBT**
  - **1X PBT** **2 X 5 min, rotating at room temperature**
- **Store in PBT at 4°C.**

Note: In the experiments of all embryos shown in this manuscript (figures 15-17), it was noticed during the fixation that a pipet was not calibrated and was ejecting ~50% of the indicated volume. The glutaraldehyde concentration was immediately adjusted to the approximate desired concentration. However, because the proteinase K permeabilization had also been performed at approximately half the desired concentration, the E11.5 and E12.5 embryos (not E10.5) were further permeabilized in 20μg/ml proteinase K (after rinsing off the fixative) for 13 minutes and 18 minutes respectively. In the last minutes of the second permeabilization, bubbles still present in the membranes around the heads of the embryos were pierced with a sterile 27½ gauge needle. The glycine rinses and fixation steps were repeated and then the experiment proceeded according to protocol.

α-Digoxigenin (DIG) antibody pre-absorption:

- **add a pinch of mouse acetone powder (Hogan 1994) to 1ml TBT + 2mM levamisole.**
- **incubate 30 min at 68°C.**
- **pellet, and transfer pellet to a fresh (15ml) tube containing a 1:1000 dilution of alkaline phosphatase conjugated α-DIG antibody (Roche, Inc.) (in TBT with 1% heat inactivated mouse serum, 2% blocking reagent).** Blocking reagent (Boeringer Mannheim, Inc.) should be prepared according to supplier's specifications.
- **rotate 1h at 4°C.**
- **pellet 10 min at 4°C, accelerating and decelerating centrifuge slowly. Transfer supernatant to a fresh tube.**
- To prepare antibody mix 1:2000 for whole mount staining, the entire pre-absorbed mixture (1:1000) is used and diluted in TBT to a final concentration of 1:2000 with 1% mouse serum (the pre-absorbed mixture contains 50% of the required serum), 2mM levamisole, 1% blocking reagent (already in the pre-absorbed mixture).

Appendix 2. Detailed protocol for *in situ* hybridization of sections

DAY 1:

- **Dewax:** **100% toluene 2 X 5 min**
  - Use full coplin jars and shake slowly to ensure complete dewaxing. Leave longer if there is wax residual on slides following second wash.
- **Rehydrate:**

100%	**EtOH**	**2 min each, shaking slowly**
95%	**EtOH**	
70%	**EtOH**	
DEPC	**water**	
DEPC	**PBS**	**2 X 5 min, shaking slowly**

  - Solutions are poured into and decanted from a coplin jar containing slides.
  - Regular grade 100% and 95% ethanol are sufficient if diluted using DEPC treated water. Solutions prepared as 50ml volumes will fill coplin jars.
- **Post-fix:** **4% formaldehyde** **10 min, shaking slowly**
  - Dilute concentrated formaldehyde (37%) in DEPC treated PBS immediately before use.
- warm TE (50ml) for permeabilization step to ~55°C and warm hybridization oven ("in slide out") to 60°C.
- **RNase Treatment:** **0.1% active DEPC in PBS** **2 X 15 min, shaking slowly**
  - Add DEPC to PBS immediately before use.
- **Permeabilization:** **20µg/ml proteinase K** **20 sec at 37°C, shaking slowly**
  - Proteinase K diluted to 20µg/ml from a 10mg/ml stock in TE (20mM Tris pH7.5, 5mM EDTA). Remove TE from 55°C ~2 min before use and add proteinase K at this time.
  - NOTE: Tris solutions **cannot** be DEPC treated.
  - Transfer slides to an empty coplin jar, pour in proteinase K solution, and quickly decant after 20 sec. Pour in the 5X SSC immediately.
- **Equilibrate:** **5X SSC (DEPC)** **2 X 10 min**
- **Pre-hybridization:** **Hybridization solution** **~2h at hybridization temperature**
  Hybridization solution: 50% formamide, 5X SSC, 500µg/ml salmon or herring sperm DNA (prepared using deionized formamide and DEPC treated solutions)
  - Space small pieces of Whatman paper containing 250µl of 25% formamide in slide holder to moisten chamber.
  - Remove slides from coplin jar, apply ~120µl hybridization solution over tissue, and slowly cover with a piece of parafilm.
- **Denature Probe:** **5 min at 80°C**
  - Thaw probes on ice, dilute to desired concentration with hybridization solution, and heat in temperature block. Remove slide holder from oven and add probe to slides 1 at a time. Lift off parafilm with SDS rinsed tweezers, tilt slide to decant, add probe directly to sections from temperature block, immediately return parafilm, and place in slide chamber.
- **Hybridization:** **overnight at desired temperature**
  optimal temperature 45-65°C (depending on probe)

DAY 2

- **Wash:**

2X SSC	**30 min, shaking slowly**
2X SSC	**60 min, shaking slowly at hybridization temperature**
0.1X SSC	**60 min, shaking slowly at hybridization temperature**

  - After hybridization, remove parafilm, and place slides in a coplin jar for washes. Wash solutions should be pre-warmed as required.
- **Equilibrate:** **Buffer 1** **2 X 5 min**
  Buffer 1: 10mM Tris pH7.5, 150mM NaCl
- **DIG probe detection: 1:2000 α-DIG antibody** **2h at room temperature**
  - Dilute alkaline phosphatase conjugated α-Digoxigenin (DIG) antibody (Roche, Inc.) in buffer 1 supplemented with 1% blocking reagent (Roche, Inc.). Prepare a 50ml re-usable stock supplemented with 0.02% sodium azide and store at 4°C.

- **Wash:** **Buffer 1** **2 X 15 min, shaking slowly at room temperature**
- **Equilibrate:** **Buffer 2** **5 min, shaking slowly at room temperature**

Buffer 2: 100mM Tris pH9.5, 100mM NaCl, 50mM $MgCl_2$

- **Colour development:** **NBT/BCIP development mix** **37°C, room temperature, or 4°C**

Development mix: 397µl Buffer 2, 1.8µl NBT, 1.4µl BCIP (Promega, Inc.)

- Place slides in slide holder, circle with hydrophobic pen, and add 400µl fresh development mix. Place Whatman sheet dampened with sterile water in bottom of cassette to moisten chamber. Start development at 37°C. If reaction is slow, leave overnight at room temperature or 4°C and continue development the next day until optimal.

DAY 3

- **Stop:** **TE pH8.0** **2 X 5 min (or overnight)**
- Once optimal, place slides in coplin jar containing TE. Slides should be left in TE until all are developed.

- **Counterstain:**	**0.5% Methyl green**	**2 min**
	distilled water	**3 X 5 sec**
- **Dehydrate:**	**35% EtOH**	**2 min each, shaking slowly**
	50% EtOH	
	75% EtOH	
	95% EtOH	
	100% EtOH	
	100% toluene	**2 X 2 min, shaking slowly**

- **Mount:** **apply DPX (BDH Laboratory Supplies, Inc.) mounting media and ease on coverslip.**

Appendix 3. Nucleotide sequences of *in situ* hybrization probes

The sequences shown correspond to the cDNA inserts used for *in vitro* transcription of antisense or sense probes. The 5' end of each transcript is shown to the left. UNC119 probes were transcribed directly from mouse IMAGE clone 3484396 (gi:12805066) which contains the full-length UNC119 cDNA cloned into vector pCMV-Sport6. UNC119B probes were transcribed from a derivative of mouse IMAGE clone 3588380 (gi:14250370) named UNC119BΔ. In UNC119BΔ, the 3146 base insert (also cloned into pCMV-Sport6) of clone 3588380 was truncated to its 5' 935 base pairs by removal of an internal XbaI fragment. The UNC119BΔ insert contains the ORF and the proximal region of the UNC119B 3' UTR. The *Pax1* probe was transcribed from a PCR amplified fragment based cloned into the pGEM®-T vector (Promega, Inc.). The amplified fragment corresponds to base pairs 2017-2560 located in the 3' UTR of the 2619 base full-length mouse *Pax1* cDNA (gi:200223). For all constructs, use of the T7 RNA polymerase generates antisense transcripts and the use of the SP6 RNA polymerase yields the sense probes.

A. UNC119 (1319 bp)

gcggcggcggtggcagcgggcgaggatctgcggccccgcgaggccatgaaggtgaagaaaggcggcggcgggaccgggtcgggggcggagccc
gttccaggggcctcgaaccggagcgcggagcccacgcgggagcctggggcggaagctgagtccgggtccgagtcggagccggagccaggcccgg
ggcccaggctggggccgctgcagggcaagcagcccatcgggccggaggatgtgctggggctgcagcggatcacgggtgactacctgtgctcccctga
ggaaaatatctacaagattgacttcgtcaggttcaagatccgtgacatggactcagggactgtccttttgaaatcaagaagccccctgtttcggaacggttgc
ccatcaaccggcgggacctggaccccaatgcagggcgctttgttcgctaccagttcacacctgccttcctccgcctaaggcaggtgggagccacggtgga
gttcacagtgggagacaagccggtcaacaacttccgcatgatcgagaggcactactttcgaaaccagctcctcaaaagcttcgacttccactttggcttctgc
atccccagcagcaagaacacctgtgagcacatctatgacttccctcctctctccgaggagctaatcagtgagatgattcgtcacccgtatgagacccagtctg
acagcttctacttcgtggatgaccggctggtgatgcataacaaagcagactattcctacagtgggacaccctgacccctaatgctgcccctgaggctctggc
ctagaccctgtgctgtaacctctccgacatcatctccacttggggatggatccggtggctctggaccctgagtcagtctctgagaggaaaggtgtccagcatc
cctaggcaaacctctggagctagtgggaggggcagcaggggggaggctggctgtccccatgttcaggaaggcctcctgcaagaggaggactcctggatg
cctcgctaggtccactgggcctgagtttcagagtagtgccccccttcttgaccgtgactagatcaggttagggaccatttcctgtcccttgcccaccacctccag
gctatttatagttgtcggttcacagacaaagatccacagtgggtaccctgctctgagccctggcctcctccatcaggggccaagaaactacccgaggtgtga
gaggggcgggtccagatcgcagcctgtaagacctggacaaactgtatatagttttcaataaacctttctgttcttgggaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaa

B. UNC119B (935 bp)

gccacggtagagttcactgtgggtgacaggccagtgacaggtttccggatgatcgagcggcactacttccgggagcggctgctgaagacctttgactttgat
ttcggcttctgcatccctagcagcagaaacacgtgtgagcacatctatgagttcccgcagctctccgaggacgtcattcgtctgatgattgaaaatccctatga
gacgcgctccgacagcttctactttgttgacaacaagctggtgatgcacaacaaggcggactatgcctataacggaggccagtaactgctgcgggagtgga
cagggaggtcctctgccgccaccacgagcaccgagcaggaattgcagctgccgcccgtcaacacccacctcgaacctggtccaggaagctgcagctgg
ggcggcggcggggttttgtgctccaaggaggtgccagccagcagtttctcttcccagtgtttgttccctgccccctagttcacagaggtattatagtgaagtaa
aagattagggtaaggttcccgggatccagataaaagattatttctgtatagtcttagctgtgcactgcttatcccgacaagcctgactgtccaggcacagctttg
tctgccagaaggcttgtgctgggctgggctgcaggttcccagccccgcaggttcctgggtgctttgggcagctgagagcagctcacttctcccccgtcccg
tgccactggcacagtctgtcctcgcctatgcagttaggccaactggaagaaacttgaagagggactggggtaggcagagcccgggctaagccactgggtt
gggaggtggcaccagagtcaaagataaggttttgaggggcttcttactgcctttgagaagccacagccaggcctgggggaccacggtctcagcctggt
ggcatcttctgttcttgct

C. Pax1 (544 bp)

cttggattgctgcccctacccaagaatattgtttggaaaccttttgtttttctgtgcccagtgactaggctggctttgtttctacgttaatgtcacccaaaagaagac
aggacaggagaaagagatgctgggactggtgttgggactggtggtgtttttagtaattaaagactaggcttatctctcagccagatatgcccaacactccttaa
tgtacctccagagccatcagcatggtttcgggtgtttggggctcagtacatttgttaattttgaagaacaaactgcttcagggtggtgaggttcattttggtttctg
gtttgcaaagaattggggtgggcaaaaacccaggaacccagaaggcttcttctggagctaggcaagagggctattaaatgatttcaagttaaggcgcactg
gtggtggggactctgagcataattcaattgttgtagtcataaacatattactggtgagagtctgacttcacctgttaaacatggctgcctttgaataaaggaaagt
ctgaaactcgccagcttg

Appendix 4. Detailed protocol for immuno-histochemical staining

- Begin immediately following fixations (described in Materials and Methods). Circle around sections with a hydrophobic pen to prevent solution dispersion. Tissue **must** remain moist throughout protocol. Volumes of solutions applied to sections should generously cover tissues (typically ~100-300µl) and be quickly and efficiently removed by aspiration.

- **Wash:** **PBT 1 X 5 min**
  PBS 2 X 5 min
- **Peroxidase attenuation:** **3% H_2O_2 diluted in methanol 1 X 15 min**
  - Re-circle sections as required.
- **Wash:** **PBT 1 X 10 min**
  PBS 2 X 10 min
- **Block:** **Blocking solution with Avidin D overnight at 4°C**
  Blocking solution: 6% goat serum, 2.4% protease-free BSA in PBT
  - Solution stored at –20°C and centrifuged at 2500rpm for 5 min prior to use.
  - Add Avidin D (Vector laboratories, Inc.) to blocking solution before use at 4 drops/ml.
  - Place slides flat in slide box containing ~1cm distilled water and seal to prevent drying.
- **Wash:** **PBS 2 X 5 min**
- **Primary Antibody:** **Dilute in blocking solution with Biotin overnight at 4°C**
  or 1-3h at room temperature
  - Add Biotin (Vector laboratories, Inc.) to blocking solution before use at 4 drops/ml.
- **Wash:** **PBT 2 X 1 min**
  1 X 15 min
  PBS 2 X 10 min
- **Secondary Antibody:** **1:200 dilution α-rabbit IgG in blocking solution**
  60min at room temperature
  - α-rabbit IgG (Vector Laboratories, Inc.) is raised in goat.
- Prepare Vectastain ABC reagent (see below) and **allow to stand** 30 min before use.
- **Wash:** **PBT 1 X 10 min**
  PBS 2 X 10 min
- **Signal amplification:** **Vectastain Elite ABC reagent 30-60 min at room temperature**
  - Prepare using 1 drop each of reagents A and B (Vector Laboratories, Inc.) per 2.5ml of PBS.
- **Wash:** **PBT 1 X 10 min**
  PBS 1 X 10 min
  2 X 5 min
- **Develop:** **DAB substrate at room temperature until colour optimal (1-15 min)**
  - Prepare using DAB substrate kit (Vector laboratories, Inc.).
- **Stop:** **distilled water 5 rinses**
  - Immerse slides at least 5 times in a large tub of water.
- **Dehydrate:** **50% EtOH (2 changes) immerse slides 10 times per change**
  70% EtOH
  100% EtOH (2 changes)
  100% xylene (3 changes)
- **Mount:** **Apply Entellan mounting media (Harleco, Inc.) and ease on coverslip.**
- **Store at room temperature.**

Appendix 5. Peptide summary report for a α–UNC119B immuno-precipitate

Mascot Search Results

(http://www.matrixscience.com/cgi/master_results.pl?REPTYPE=Peptide&file=../data/20020802/FAGTmzu.dat)

Search title : 2p60_a
MS data file : C:\WINNT\Profiles\micromass\DESKTOP\2p60\2p60_a.pkl
Database : NCBInr 20020718 (1034279 sequences; 327146032 residues)
Timestamp : 2 Aug 2002 at 22:28:40 GMT
Significant hits : gi|229552 albumin [Bos taurus]

Probability Based Mowse Score
Score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 49 indicate identity or extensive homology (p<0.05).

PEPTIDE SUMMARY REPORT

1. gi|229552 **Mass:** 66088 **Total score:** 132 **Peptides matched:** 2
albumin [Bos taurus]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
5	700.35	1398.69	1398.69	0.00	0	67	1	TVMENFVAFVDK
7	784.38	1566.75	1566.74	0.01	0	65	1	DAFLGSFLYEYSR

Proteins matching the same set of peptides:
gi|418694 **Mass:** 69225 **Total score:** 132 **Peptides matched:** 2
serum albumin precursor [validated] - bovine
gi|1351907 **Mass:** 69248 **Total score:** 132 **Peptides matched:** 2
Serum albumin precursor (Allergen Bos d 6)
gi|2190337 **Mass:** 69278 **Total score:** 132 **Peptides matched:** 2
(X58989) serum albumin [Bos taurus]

2. gi|457366 **Mass:** 35270 **Total score:** 43 **Peptides matched:** 2
(L29172) Ig gamma H-chain C-region [Oryctolagus cuniculus]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
4	601.31	1200.60	1201.61	-1.01	0	19	1	LSVPTSEWQR
8	545.64	1633.89	1634.88	-0.99	0	25	1	VVSTLPIAHEDWLR

3. gi|2136983 **Mass:** 20563 **Total score:** 43 **Peptides matched:** 2
Ig gamma heavy chain constant region - rabbit (fragment)

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
4	601.31	1200.60	1201.61	-1.01	0	19	1	LSVPTSEWQR
8	545.64	1633.89	1633.89	-0.01	0	24	2	VVSTLPIAHQDWLR

4. gi|19074193 **Mass:** 25042 **Total score:** 26 **Peptides matched:** 1
(NC_003231) hypothetical protein [Encephalitozoon cuniculi]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
2	437.76	873.50	873.45	0.04	0	26	1	WILSCPR

5. gi|7228413 **Mass:** 46462 **Total score:** 26 **Peptides matched:** 1
myosin VIIA-like protein [Mizuhopecten yessoensis]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
9	636.65	1906.92	1908.91	-1.99	1	26	2	DQCVIISGESGAGKTESTK

6. gi|6754392 **Mass:** 194164 **Total score:** 25 **Peptides matched:** 1
(NM_010587) intersectin (SH3 domain protein 1A); Eh domain, SH3 domain regulator of endocytosis 1; intersection (SH3 domain protein 1A) [Mus musculus]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
1	428.75	855.48	857.45	-1.97	1	25	1	RLAMDPR

Proteins matching the same set of peptides:
gi|20071578 **Mass:** 49446 **Total score:** 25 **Peptides matched:** 1
(BC026532) Similar to intersectin 2 [Mus musculus]
gi|20560920 **Mass:** 195300 **Total score:** 25 **Peptides matched:** 1
(XM_048621) similar to intersectin long isoform [Homo sapiens]
gi|20896223 **Mass:** 193313 **Total score:** 25 **Peptides matched:** 1
(XM_122981) intersectin (SH3 domain protein 1A) [Mus musculus]

Appendix 6. α-UNC119 co-immuno-precipitate protein summary report: Ig Heavy Chain

Mascot Search Results

(http://www.matrixscience.com/cgi/master_results.pl?file=../data/20020227/FaEmoic.dat&REPTYPE=protein)

Search title : Chantal6
MS data file : C:\automsms\P-02-0034°C .pkl
Database : NCBInr 20020220 (883557 sequences; 276611002 residues)
Timestamp : 27 Feb 2002 at 21:37:36 GMT
Top Score :123 for gi|109278, Ig heavy chain V-A1 region (RVH122) - rabbit (fragment)

Probability Based Mowse Score
Score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 72 are significant (p<0.05).

PROTEIN SUMMARY REPORT

Index

	Accession	Mass	Score	Description
1.	gi|109278	5975	123	Ig heavy chain V-A1 region (RVH122) - rabbit (fragment)
2.	gi|165122	6882	122	(K01356) Ig gamma chain VDJ-region [Oryctolagus cuniculus]
3.	gi|165142	7622	121	(M21266) Ig heavy region V-D-region [Oryctolagus cuniculus]
4.	gi|4096203	10738	119	(U26485) Ig heavy chain [Oryctolagus cuniculus]
5.	gi|4096215	11194	118	(U26493) Ig heavy chain [Oryctolagus cuniculus]
6.	gi|4096197	11825	118	(U26482) Ig heavy chain [Oryctolagus cuniculus]
7.	gi|9799129	11930	118	(AF264463) IgM heavy chain VDJ region [Oryctolagus cuniculus]
8.	gi|10179993	12172	118	(AF294963) anti-guinea pig C3 single chain variable fragment antibody heavy chain [synthetic construct]
9.	gi|165132	12183	118	(M21261) Ig heavy chain V-region precursor [Oryctolagus cuniculus]
10.	gi|9799131	12103	118	(AF264464) IgM heavy chain VDJ region [Oryctolagus cuniculus]
11.	gi|482289	12315	118	Ig heavy chain precursor VH1-A1 region - rabbit (fragment)
12.	gi|4096199	12137	118	(U26483) Ig heavy chain [Oryctolagus cuniculus]
13.	gi|9799153	12200	118	(AF264475) IgM heavy chain VDJ region [Oryctolagus cuniculus]
14.	gi|4261933	12415	118	(S76746) Unknown [Oryctolagus cuniculus]
15.	gi|2707675	12774	117	(AF029941) IgM heavy chain VDJ region [Oryctolagus cuniculus]
16.	gi|2707719	12671	117	(AF029963) IgM heavy chain VDJ region [Oryctolagus cuniculus]
17.	gi|9799165	12708	117	(AF264481) IgM heavy chain VDJ region [Oryctolagus cuniculus]
18.	gi|2707669	12659	117	(AF029938) IgM heavy chain VDJ region [Oryctolagus cuniculus]
19.	gi|9799155	12839	117	(AF264476) IgM heavy chain VDJ region [Oryctolagus cuniculus]
20.	gi|9799167	12933	117	(AF264482) IgM heavy chain VDJ region [Oryctolagus cuniculus]

Results List

1. gi|109278 **Mass:** 5975 **Score:** 123
Ig heavy chain V-A1 region (RVH122) - rabbit (fragment)

Observed	Mr(expt)	Mr(calc)	Delta	Start	End	Miss	Ions	Peptide
595.34	594.33	594.34	-0.01	28	32	0	27	FTISK
917.42	1832.82	1832.83	-0.00	41	56	0	76	ITSPTTEDTATYFCAR
917.42	1832.83	1832.83	0.00	41	56	0	26	ITSPTTEDTATYFCAR

No match to: 505.26, 619.34, 523.28, 523.28, 523.29, 523.29, 523.29, 595.34, 599.77, 605.34, 671.84, 691.33, 731.85, 734.37, 734.39, 734.40, 755.93, 824.90, 854.42, 910.41, 912.99, 917.42, 924.42, 932.45, 934.42, 1028.45, 1109.48, 1109.48, 1178.99, 1190.50, 1211.01, 1211.01, 1211.01, 1211.02, 1221.49, 1280.53, 1284.04, 1292.04, 1312.55, 1361.55

Appendix 7. α-UNC119 co-immuno-precipitate protein summary report: A-X actin

Mascot Search Results

(http://www.matrixscience.com/cgi/master_results.pl?file=../data/20020206/FamTrbu.dat&REPTYPE=protein)

Search title : Chantel7
MS data file : C:\automsms\Chantel7.pkl
Database : NCBInr 20020118 (851328 sequences; 268254945 residues)
Timestamp : 6 Feb 2002 at 18:13:35 GMT
Top Score : 188 for gi|309090, (J04181) A-X actin [Mus musculus]

Probability Based Mowse Score
Score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 72 are significant (p<0.05).

PROTEIN SUMMARY REPORT

Index

	Accession	Mass	Score	Description
1.	**gi|309090**	42009	188	(J04181) A-X actin [Mus musculus]
2.	**gi|6716561**	41990	188	(AF168615) beta-actin [Rivulus marmoratus]
3.	**gi|49868**	39446	187	(X03765) put. beta-actin (aa 27-375) [Mus musculus]
4.	**gi|15277503**	40536	186	(BC012854) Unknown (protein for MGC:9832) [Homo sapiens]
5.	**gi|14250401**	41321	186	(BC008633) actin, beta [Homo sapiens]
6.	**gi|809561**	41335	186	(X13055) gamma-actin [Mus musculus]
7.	**gi|13278939**	41506	186	(BC004223) Unknown (protein for IMAGE:3533309) [Homo sapiens]
8.	**gi|224306**	41798	186	actin [Strongylocentrotus purpuratus]
9.	**gi|71621**	41921	186	actin beta - bovine (tentative sequence)
10.	**gi|71625**	41977	186	actin gamma - bovine (tentative sequence)
11.	**gi|3182897**	41971	186	ACTIN, CYTOPLASMIC
12.	**gi|231506**	42071	186	ACTIN, CYTOPLASMIC 1 (BETA-ACTIN)
13.	**gi|481515**	42128	186	actin beta' chain - human
14.	**gi|7546744**	42068	186	(AF135499) beta-actin [Platichthys flesus]
15.	**gi|3348131**	42082	186	(AF079161) cytoplasmic beta actin [Xenopus laevis]
16.	**gi|1703118**	42041	186	ACTIN, CYTOPLASMIC 3 (BETA-ACTIN 3)
17.	**gi|2318133**	42033	186	(AF014363) beta actin [Cricetinae gen. sp.]
18.	**gi|6230866**	41957	186	(AB034209) cytoplasmic actin [Oikopleura longicauda]
19.	**gi|1351867**	42053	186	ACTIN, CYTOPLASMIC 1 (BETA-ACTIN)
20.	**gi|16304154**	42034	186	(AF421789) beta actin [Sigmodon hispidus]

Results List

1. gi|309090 **Mass:** 42009 **Score:** 188
(J04181) A-X actin [Mus musculus]

Observed	Mr(expt)	Mr(calc)	Delta	Start	End	Miss	Ions	Peptide
566.75	1131.48	1131.52	-0.04	197	206	0	22	GYSFTTTAER
581.30	1160.59	1160.61	-0.02	316	326	0	27	EITALAPSTMK
599.75	1197.49	1197.51	-0.02	51	61	0	43	DSYVGDEAQSK
599.78	1197.54	1197.51	0.02	51	61	0	----	DSYVGDEAQSK
758.83	1515.65	1515.70	-0.04	360	372	0	28	QEYDESGPSIVHR
895.92	1789.82	1789.88	-0.06	239	254	0	25	SYELPDGQVITIGNER

No match to: 594.27, 629.23, 681.24, 788.21, 911.38, 990.41, 507.73, 534.18, 557.76, 589.29, 589.30, 602.27, 602.27, 602.27, 750.32, 750.33, 926.39

2. gi|6716561 **Mass:** 41990 **Score:** 188
(AF168615) beta-actin [Rivulus marmoratus]

Observed	Mr(expt)	Mr(calc)	Delta	Start	End	Miss	Ions	Peptide
566.75	1131.48	1131.52	-0.04	197	206	0	22	GYSFTTTAER
581.30	1160.59	1160.61	-0.02	316	326	0	27	EITALAPSTMK
599.75	1197.49	1197.51	-0.02	51	61	0	43	DSYVGDEAQSK
599.78	1197.54	1197.51	0.02	51	61	0	----	DSYVGDEAQSK
758.83	1515.65	1515.70	-0.04	360	372	0	28	QEYDESGPSIVHR
895.92	1789.82	1789.88	-0.06	239	254	0	25	SYELPDGQVITIGNER

No match to: 594.27, 629.23, 681.24, 788.21, 911.38, 990.41, 507.73, 534.18, 557.76, 589.29, 589.30, 602.27, 602.27, 602.27, 750.32, 750.33, 926.39

3. gi|49868 **Mass:** 39446 **Score:** 187
(X03765) put. beta-actin (aa 27-375) [Mus musculus]

Observed	Mr(expt)	Mr(calc)	Delta	Start	End	Miss	Ions	Peptide
566.75	1131.48	1131.52	-0.04	171	180	0	22	GYSFTTTAER
581.30	1160.59	1160.61	-0.02	290	300	0	27	EITALAPSTMK
599.75	1197.49	1197.51	-0.02	25	35	0	43	DSYVGDEAQSK
599.78	1197.54	1197.51	0.02	25	35	0	----	DSYVGDEAQSK
758.83	1515.65	1515.70	-0.04	334	346	0	28	QEYDESGPSIVHR
895.92	1789.82	1789.88	-0.06	213	228	0	25	SYELPDGQVITIGNER

No match to: 594.27, 629.23, 681.24, 788.21, 911.38, 990.41, 507.73, 534.18, 557.76, 589.29, 589.30, 602.27, 602.27, 602.27, 750.32, 750.33, 926.39

Appendix 8. α-UNC119 co-immuno-precipitate protein summary report: EpoR

Mascot Search Results

(http://www.matrixscience.com/cgi/master_results.pl?file=../data/20020206/FamTSns.dat&REPTYPE=protein)

MS data file : C:\automsms\Chantel14.pkl
Database : NCBInr 20020118 (851328 sequences; 268254945 residues)
Timestamp : 6 Feb 2002 at 20:03:43 GMT
Top Score : 31 for gi|193200, (M62360) erythropoietin receptor protein [Mus musculus]

Probability Based Mowse Score
Score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 72 are significant (p<0.05). Protein Summary Report

PROTEIN SUMMARY REPORT

Index

	Accession	Mass	Score	Description
1.	gi|193200	2674	31	(M62360) erythropoietin receptor protein [Mus musculus]
2.	gi|11497036	4458	28	(NC_001903) B. burgdorferi predicted coding region BBB20 [Borrelia burgdorferi]
3.	gi|7524649	6558	27	(NC_001631) ORF56a [Pinus thunbergii]
4.	gi|9827841	6692	27	(AL160493) hypothetical protein LM26.312 [Leishmania major]
5.	gi|9625014	19883	27	(NM_019610) hypothetical protein 669 [Homo sapiens]
6.	gi|17546225	6799	27	(NC_003295) PROBABLE TRANSMEMBRANE PROTEIN [Ralstonia solanacearum]
7.	gi|4826622	8110	27	(AF117644) kinesin tail 911 [Expression vector pH911]
8.	gi|6049602	7636	26	(AF175706) nucleocapsid protein [Hantavirus HTN/Far East/7732]
9.	gi|17402626	7590	26	(AF191686) odorant receptor [Salmo salar]
10.	gi|17506895	25152	26	(NM_060079) Yeast hypothetical protein YG15 like [Caenorhabditis elegans]
11.	gi|2707290	6714	26	(AF036318) cyclin [Skeletonema costatum]
12.	gi|17228150	7323	26	(NC_003272) hypothetical protein [Nostoc sp. PCC 7120]
13.	gi|14602300	6611	26	(NC_002816) ORF63 BRO [Cydia pomonella granulovirus]
14.	gi|17864913	8033	26	(AF447808) unknown [Haemophilus influenzae biotype aegyptius]
15.	gi|950409	7207	26	(L46723) galactose-1-phosphate uridyl transferase [Homo sapiens]
16.	gi|16271831	8262	26	(NC_003157) hypothetical protein-phage associated [temperate phage PhiNIH1.1]
17.	gi|12654043	8316	26	(BC000825) Unknown (protein for IMAGE:3454279) [Homo sapiens]
18.	gi|9294173	8518	26	(AB026647) gb|AAD12021.1~gene_id:MJL12.11~similar to unknown protein [Arabidopsis thaliana]
19.	gi|17647325	9387	26	(NM_078984) Developmental embryonic B [Drosophila melanogaster]
20.	gi|15081491	9498	26	(NC_003042) probable transposase [Clostridium perfringens]

Results List

1. gi|193200 **Mass:** 2674 **Score:** 31
(M62360) erythropoietin receptor protein [Mus musculus]

Observed	Mr(expt)	Mr(calc)	Delta	Start	End	Miss	Ions	Peptide
678.15	2031.43	2031.15	0.28	6	24	1	----	VPLWPRVGPLCLLLAGAAW